

Projektbericht

der Projektgruppe 1.5 mit dem Thema

„Mehr Bewegung an Schulen“



Projektbetreuung & Prüferin: Rabea Wienholt

Abgabedatum: 01.07.2022

Autorinnen:

Universität: Technische Universität Dortmund

Fakultät: Rehabilitationswissenschaften

Studiengang: Bachelor of Arts Rehabilitationspädagogik

Abstract

Eine Studie der Weltgesundheitsorganisation (WHO) (2019) hat ergeben, dass Jugendliche zu wenig körperlich aktiv sind und mehr als 80% der Kinder und Jugendlichen sich weniger als eine Stunde am Tag bewegen. Dabei liegt die Empfehlung der Nationalen Bewegungsempfehlung bei etwa 90 Minuten. Die aktuelle Covid-19-Pandemie erschwerte, sowohl durch Schließung der Schulen als auch der Freizeiteinrichtungen das Erreichen dieser Zeiten enorm. Um dem Bewegungsmangel an Schulen ein Stückweit entgegenzuwirken, setzt das Projekt mit bewegten Pausen innerhalb des Unterrichts an. Der vorliegende Forschungsbericht geht demgemäß der Forschungsfrage nach, inwiefern sich das psychomotorische Beobachtungsinstrument MotorikPlus von Zimmer (2021) auf die bestimmten Altersspannen adaptieren lässt. Ziel der Projektgruppe ist es, den schon bestehenden psychomotorischen Test von Zimmer auf andere Altersspannen zu adaptieren, um somit einen Einsatz in der Grundschule sowie der Sekundarstufe I gewährleisten zu können. Für die Beantwortung dieser Forschungsfrage wurden die Daten mithilfe des Auswertungsprogramm RStudio aufgearbeitet und analysiert, woraufhin weitere Modifikationen am Manual vorgenommen werden konnten. Außerdem wurde das Erhebungsinstrument mit einer qualitativen schriftlichen Befragung, um einen tiefergehenden Einblick in die Lehrer*innen Perspektive zu erlangen und dementsprechend mehr Informationen über die durchgeführten Interventionen innerhalb der Klassen zu sammeln, ergänzt.

Inhaltsverzeichnis

1. Einleitung	5
2. Fachliche Einführung	6
2.1 Kooperationsschulen	6
2.2 Setting der Erhebungen	8
2.3 Rahmung durch den MotorikPlus	9
3. Projektmethodik und -design	10
3.1 Forschungsfrage	10
3.2 Projektziele	11
3.3 Entwicklung des Erhebungsinstrumente	11
3.4 Stichprobe	12
3.5 Untersuchungsdesign	13
3.5.1 Qualitatives Forschungsdesign	13
3.5.2 Quantitatives Forschungsdesign	15
3.5.3 Ablauf der Datenerhebungen	17
4. Darstellung der quantitativen Forschungsergebnisse	19
4.1 Objektivität	19
4.2 Validität	23
4.3 Reliabilität	23
4.3.1 Itemschwierigkeiten	24
4.3.2 Cronbachs Alpha	27
4.3.3 Trennschärfe Koeffizienten	29
5. Überarbeitung Manual	32
5.1 Ü6 Bereich	32
5.2 Ü9 Bereich	37
6. Diskussion	42
6.1 Beantwortung Forschungsfrage	42
6.2 Bewertung der Projektziele	42
6.3 Kritische Reflexion der Projektergebnisse	44
6.4 Limitation der Forschung	45
6.5 Forschungsbedarf	45
6.6 Kritische Reflexion Projektverlauf & Projektorganisation	46
6.7 Rückblick	47
7. Fazit	50
8. Literaturverzeichnis	51

9. <i>Abbildungsverzeichnis</i>	53
10. <i>Anhang</i>	54
11. <i>Eidesstaatliche Versicherung</i>	62

1. Einleitung

Die Lebensführung in den westlichen Industrienationen, als auch die gesellschaftlichen Wandlungen, wie die Corona Pandemie, führen zu einem bewegungsinktiven Lebensstil bei Kindern. Vor dem Hintergrund, dass Bewegungsmangel ein äußerst signifikantes und ernstes Kernproblem der Gesellschaft darstellt, sowie die Abnahme an körperlicher Aktivität bei Grundschulkindern hervorruft, wurde von der Projektgruppe 1.5, ein quantitatives Beobachtungsinstrument durch die Adaption des MotorikPlus von Zimmer für die Altersspanne Ü6 und Ü9 entwickelt. In der Vorliegenden Arbeit wird zunehmend das Erhebungsinstrument sowie Beobachtungsverfahren, für die Messung des Einflusses von bewegten Pausen im Unterricht auf die psychomotorische Entwicklung von Kindern und Jugendlichen, dargestellt.

Das Beobachtungsverfahren wird auf der Grundlage von wissenschaftlichen Methoden, welche sich bestimmter zielgerichteter Techniken, d.h. einem System von Regeln, Vorgehensweisen, Prinzipien und Verfahren bedienen, behandelt. Daher lässt sich auch der zweite Teil des Projektberichts, in vier Themenblöcke unterteilen. Der erste Teil des Berichts befasst sich mit der fachlichen Einführung und thematisiert unter anderem die teilnehmenden Kooperationsschulen, das Setting der Erhebung sowie die Rahmung durch den MotorikPlus. Im Fokus des zweiten Teils steht die Projektmethodik sowie das qualitative- und quantitative Forschungsdesign. Die Stichprobe wurde in drei Interventionsarten unterteilt: der Bewegungsklasse, Leseklasse, sowie der Kontrollklasse. Des Weiteren widmet sich der dritte Themenblock der Darstellung der quantitativen Forschungsergebnisse unter Betrachtung der Gütekriterien. Darauf aufbauend werden die aufgeführten Ergebnisse zusammenfassend diskutiert und mit einer kritischen Reflexion des Projektverlaufs und der Projektorganisation sowie einem Rückblick abgeschlossen.

2. Fachliche Einführung

Die fachliche Einführung gibt einen Einblick in die praktische Arbeit des Projektes. Zunächst werden die Kooperationsschulen erwähnt und es wird drauf eingegangen in welchem Setting die Untersuchungen stattgefunden haben. Abschließend wird eine kurze Rückbindung an den MotorikPlus erfolgen, welcher als Grundlage des Projektes dient.

2.1 Kooperationsschulen

Innerhalb des Projektes wurden wissenschaftliche Erhebungen mittels des Testinstrument in verschiedenen Schulen vorgenommen. Diese fanden in einem Zeitraum von sieben Monaten statt. Aufgrund des immer stärker werdenden Bewegungsmangel, welchem das Projekt entgegenwirken möchte, wurde das Interesse der Schulen, durch ein zugeschicktes Exposee mit Informationen über das Projekt, geweckt (BMFG, 2020). Die wissenschaftlichen Erhebungen fanden an der Schule an der Bergmannsglückstraße in Bottrop und Gelsenkirchen, sowie an der Fährmannschule in Recklinghausen statt. Weitere Schüler*innen von anderen Schulen konnten aufgrund der Covid-19-Pandemie nicht untersucht werden. Die Schule an der Bergmannsglückstraße gibt es, wie oben schon genannt, in Bottrop sowie in Gelsenkirchen, da die Schule an zwei Standorten verankert ist. Beide Standorte nahmen an der Studie teil. Bei der zuvor genannten Schule handelt es sich um eine sonderpädagogische Schule mit dem Förderschwerpunkt emotionale und soziale Entwicklung. Es nahmen insgesamt drei Klassen der Schule an der Studie teil. Darunter befand sich die Kontrollklasse im Standort Bottrop neben der Lese- sowie Bewegungsklasse in Gelsenkirchen. Die Schüler*innen in den Klassen waren in dem Altersbereich Ü9 (10-15 Jahre). Für die Kontrollklasse in Bottrop war Frau W. verantwortlich. Die Kooperation mit der vorgenannten Klasse verlief unproblematisch. Für die Leseklasse in Gelsenkirchen war Herr H. verantwortlich, auch hier verlief die Kooperation ohne Probleme. Als Leseintervention wurden diverse Bücher gelesen, wie „Das fliegende Klassenzimmer“, „Drei Fragezeichen Romane“ und „100-Gefahren-Geschichten“. Wenn möglich wurde täglich ein Kapitel für circa 15 bis 20 Minuten gelesen (Anhang A, Z. 11ff). Herr H. zeigte großes Interesse und war sehr bemüht die Intervention ordnungsgemäß und regelmäßig durchzuführen. Er teilte uns mit was in der Klasse

vorgelesen wurde und in welchem Umfang er die Leseintervention einsetzte. Frau B. war die verantwortliche Lehrerin für die Bewegungsklasse. Als Bewegungsintervention wurden angeleitete Rennspiele vor Schulbeginn in einem zeitlichen Umfang von circa 15 Minuten durchgeführt sowie ein abwechselndes Angebot, wie beispielsweise verlängerte Bewegungspausen mit Ballspielangeboten in einem zeitlichen Umfang von 2-mal circa 10 Minuten. Außerdem wurden kurze Bewegungsangebote, wie beispielsweise Yoga oder Stretching während des Unterrichts und bei Phasenwechseln angeboten und angeleitet, in einem zeitlichen Umfang von circa 15 Minuten (Anhang A, Z.39ff). Hier gab es vereinzelte Probleme, da die Kooperation zum Teil etwas schwieriger verlief. Beispielsweise wurde im Verlauf des Prozesses keine Auskunft erteilt, inwiefern die Bewegungsintervention innerhalb des Projektes umgesetzt wurden. Erst bei einer schriftlichen Befragung wurden diese Informationen nachgetragen. Allerdings wurde durch diese ebenso ersichtlich, dass sie sich mehr Anleitung oder auch Begleitung bei der Gestaltung der Interventionen gewünscht hätte (Anhang A, Z. 47f). Zusätzlich dazu herrschte eine hohe Fluktuation innerhalb der Klasse, weshalb sich das Klassengefüge stetig veränderte und somit immer wieder neue Kinder in das Projekt eingeführt werden mussten (Anhang A, Z. 36f). Die Fährmannschule, mit dem Förderschwerpunkt Sprache im Primarbereich, hat ihren Standpunkt in Recklinghausen. Die Schüler*innen lagen im Altersbereich Ü6 (6-10 Jahre). In Recklinghausen gab es eine Lese- und Bewegungsklasse, sowie eine Kontrollklasse. Bei der Bewegungsklasse wurde als Intervention zwei Mal täglich sportliche Übungen durchgeführt (Anhang A, Z.66). Die hauptverantwortliche Bezugsperson war Frau H. Die Kooperation verlief bei der ersten Datenerhebung unproblematisch und die Lehrer*innen gaben großes Interesse kund. Jedoch kam es zu keiner zweiten und dritten Datenerhebung, da nach kontinuierlichem Kontaktversuch keiner zu erreichen war und somit keine Termine ausgemacht werden konnten.

Aufgrund der verschiedenen Förderschwerpunkte kam es zu unterschiedlichen Erschwernissen. In Recklinghausen hatten die Schüler*innen, aufgrund des Förderschwerpunkts Sprache, Schwierigkeiten die Aufgabe zu verstehen, was dazu führte, dass sie Aufgaben nicht absolvieren konnten, weil die Aufgabenstellung nicht in leichter Sprache formuliert war. Außerdem kam es vereinzelt vor, dass sie Buchstaben nicht lesen konnten. Auch hier konnte beispielsweise die Aufgabe „Rasterbild

erkennen“ aufgrund des Förderschwerpunktes nicht von allen Kindern richtig gelöst werden (Anhang C, Z. 31f). In Gelsenkirchen und Bottrop mit dem Förderschwerpunkt emotionale und soziale Entwicklung kam es zu anderen Schwierigkeiten. Bei den Erhebungen haben sich die Schüler*innen oft untereinander provoziert oder auch den Drang dazu gehabt sich zu vergleichen und besser als der/die andere zu sein. Die Atmosphäre war dadurch angespannter und mit mehr Energie geladen.

2.2 Setting der Erhebungen

Wie bereits erwähnt fanden die Datenerhebungen in verschiedenen Schulen statt, sodass auch die Settings für die Erhebungen innerhalb der Schulen variiert haben. Die erste Erhebung fand am 10.11.2021 in Gelsenkirchen, für die Schüler*innen der Bewegungs- und Leseklasse, in einem normalen Klassenraum statt. Hier wurden einzelne Stationen auch im Flur aufgebaut, da der Klassenraum nicht groß genug war (Anhang C, Z. 5ff). Die Stationen waren aufgrund von Platzmangel sehr nah aneinander aufgebaut. Die zweite und dritte Erhebung in Gelsenkirchen fanden am 20.01.2022 und 15.05.2022 in der Aula statt (Anhang C, Z. 34f). Dort war genug Platz, um alle Stationen innerhalb dieser Räumlichkeit aufzubauen und durchzuführen. Allerdings gingen mit diesem Raum einige Störvariablen einher. Zum einen war der Boden rutschig. Zum anderen war die Pausenklingel sehr laut, sodass die Testung aufgrund dessen pausiert werden musste (Anhang C, Z. 36ff). Es stand genügend Bewegungsfreiraum zur Verfügung, sodass alle Aufgaben, gerade Übungen mit sportlicher Aktivität, wie Balancieren, über ein Hindernis springen oder das Werfen auf die Zielscheibe, gut ausgeführt werden konnten. In Bottrop fanden alle drei Erhebungen im Psychomotorik Raum der Schule statt. Diese wurden am 12.11.2021, 21.01.2022 und 19.05.2022 durchgeführt. Auch hier bot der Raum genug Platz, um alle Stationen im Raum durchführen zu können (Anhang C, Z.19f). In Recklinghausen wurde die erste Erhebung am 17.11.2021 ebenfalls in der Aula durchgeführt, welche groß und geräumig war, sodass alle Stationen ohne Platzprobleme durchgeführt werden konnten (Anhang C, Z.24).

Die Datenerhebungen wurden in den verschiedenen Klassen, mit unterschiedlichen Förder- und Altersschwerpunkt durchgeführt, um möglichst viele Daten zu erfassen. Diese dienen zum einen dazu, eine repräsentative Stichprobe zu erlangen und zum

anderen Vergleiche ziehen zu können. Aufgrund von mangelnden Kooperationsschulen, im Hinblick auf die Anzahl von Personen und die Schulform, konnte innerhalb des Projektes keine repräsentative Stichprobe entstehen.

2.3 Rahmung durch den MotorikPlus

Um die theoretische Besonderheit des Projektes klarzustellen, wird folgend Bezug dazu genommen. Die aktuelle Covid-19 Pandemie reduzierte die Bewegung im Alltag, wodurch der Bewegungsmangel verstärkt wurde. Demzufolge hat das Projekt eine große gesundheitliche Relevanz und verfolgt das Ziel dem Bewegungsmangel entgegenzuwirken (BMFG, 2020). Als Grundlage hierfür dient das Beobachtungsinstrument „MotorikPlus“ von Zimmer (2021). Mithilfe des Instruments können die psychomotorischen Kompetenzen von Kindern im Alltag von Kindertageseinrichtungen beobachtet werden (Zimmer, 2021). Demzufolge umfasst die Zielgruppe aus dem MotorikPlus, Kinder aus dem U3 sowie Ü3 Bereich (1;0-6;11 Jahre). Das selbstentwickelte Messinstrument hingegen beschäftigt sich mit den Altersstufen Ü6 und Ü9, da der Bewegungsmangel nicht nur im Kindergartenalter, sondern auch im Grundschulalter sowie in der Sekundarstufe I vorhanden ist. Ebenso soll mithilfe der Adaption auf den Ü6er und Ü9er Bereich ein Einsatz des Beobachtungsinstruments in der Grundschule sowie der Sekundarstufe I gewährleistet werden. Außerdem handelt es sich bei dem MotorikPlus um einen alltagstauglichen Test, da viele Items wie beispielsweise „Das Kind kann beim Laufen die Richtung schnell ändern“ gut im Kindergartenalltag durchgeführt und integriert werden können (Zimmer, 2021). Dies wiederum entspricht nicht dem selbstentwickelten Testinstrument, da die Aufgaben kaum in den Schulalltag integriert werden können, sondern vielmehr aktive Testsituationen erforderlich sind. Aufgrund der bisher statischen Lehrform innerhalb von Schulen, beziehungsweise unterschiedlicher Lehrinrichtungen, gibt es wenig Handlungsspielraum die Items innerhalb des Schulalltags durchzuführen, um die Kompetenzen der Kinder zu beobachten.

3. Projektmethodik und -design

Im darauffolgenden Unterkapitel werden theoretische Fakten zu der Projektmethodik und dem Projektdesign erläutert. Zunächst wird ein Einblick in die Forschungsfrage und die Projektziele gegeben. Folgend wird Bezug auf die Entwicklung des Erhebungsinstruments genommen und die Stichprobe wird in ihrer Art und ihrem Umfang durchleuchtet. Diese Unterkapitel bilden das methodische Vorgehen und werden ergänzt durch das Untersuchungsdesign, welches sich in die qualitative sowie quantitative Forschung einteilen lässt.

3.1 Forschungsfrage

Der vorliegende Forschungsbericht geht der Forschungsfrage nach, inwiefern sich das psychomotorische Beobachtungsinstrument MotorikPlus von Zimmer (2021) auf andere Altersspannen adaptieren lässt. Im Folgenden werden die Daten des adaptierten Messinstruments ausgewertet, analysiert und folglich interpretiert, um diese Forschungsfrage beantworten zu können. Im Verlauf der Literaturrecherche wurde ersichtlich, dass sich der Bewegungsmangel nicht lediglich auf die von Zimmer behandelte Altersstufe begrenzt, weswegen das Ziel, den MotorikPlus auf die Altersstufen Ü6 und Ü9 zu adaptieren, gesetzt wurde. Zu Beginn des Projekts wurde sich zunächst mit der Frage „Inwiefern haben bewegte Pausen in der Schule Einfluss auf die psychomotorische Entwicklung von Kindern und Jugendlichen?“ beschäftigt. Aufgrund der Gruppengröße und dem noch nicht vorhandenen Erhebungsinstrument für die betreffende Altersspanne, wurde sich lediglich auf die Adaption fokussiert und ob das Messinstrument, das misst was es messen soll. Mit einem final fertiggestellten und zudem normierten Messinstrument für diese Altersspanne könnte diese Forschungsfrage tiefergehend angegangen werden. Das Forschungsprojekt legt die Grundlage für spätere Forschungsvorhaben und ist dementsprechend eine Pilotstudie. Ebenso dient es als explorative Voruntersuchung, die zur Theorie- und Hypothesenbildung anregt (Spektrum, n.d.).

3.2 Projektziele

Der Forschungsfrage untergeordnet wurden noch weitere Ziele im Verlauf des Projekts angestrebt. Demzufolge sollten weitere Erhebungen stattfinden, um das Manual dadurch verbessern zu können. Aus diesen werden Daten gesammelt die zum einen zur Weiterentwicklung des Manuals beitragen und zum andern die Durchführungsqualität bei den Erhebungen steigern können. Unter anderem lag es im Interesse der Projektgruppe, die psychomotorischen Kompetenzen mithilfe des adaptierten quantitativen Messinstruments von Kindern und Jugendlichen in der Grundschule sowie der Sekundarstufe I messen zu können. Demzufolge wurden die Daten mithilfe von dem Auswertungsprogramm RStudio aufgearbeitet und analysiert, woraufhin weitere Modifikationen am Manual vorgenommen werden konnten. Zusätzlich zu der quantitativen Analyse lag ein weiteres Ziel darin das Projekt qualitativ durchleuchten zu können. Hierfür wurde eine qualitative schriftliche Befragung durchgeführt, um einen tiefergehenden Einblick in die Lehrer*innen Perspektive zu erlangen und dementsprechend mehr Informationen über die durchgeführten Interventionen innerhalb der Klassen zu sammeln. Die untergeordneten Ziele sollen dabei helfen die übergeordnete Zielsetzung zu beantworten.

3.3 Entwicklung des Erhebungsinstruments

Bei der Entwicklung des selbst erstellten quantitativen Erhebungsmessinstruments wurde als Grundlage der MotorikPlus von Zimmer (2021) genutzt. Zu Beginn des Projektstudiums wurde Literatur gesichtet, um beispielsweise einen Überblick von möglichen Aufgaben für den Altersbereich Ü6 und Ü9 zu erlangen sowie sich in die Relevanz und Forschungslücke einzulesen. Auch hier wurde der MotorikPlus als Grundlage verwendet, da dieser viele motorische, sensorische sowie psychomotorische Aufgaben beinhaltet. Nachdem Aufgaben zum Teil adaptiert oder auch neu entwickelt wurden ging es an die praktische Umsetzung. Hierfür wurde das Bewegungsambulatorium (BWA) an der technischen Universität Dortmund genutzt. Es wurden einzelne Aufgaben und Instruktionen getestet, um eine erste Adaption zu erstellen. Dies erstreckte sich über mehrere Wochen und Monate, sodass eine fortlaufende Überarbeitung stattgefunden hat. Im November begann die erste Erhebungsphase, bei der deutlich wurde welche Aufgaben gut umgesetzt werden konnten und nicht

großartig verändert werden mussten. Bei anderen Aufgaben, wie beispielsweise dem „Perlen auffädeln“, wurde deutlich, dass keine Zeitbegrenzung festgelegt wurde und somit jede*r Beobachter*in anders gemessen hat. Nach den jeweiligen Erhebungsphasen wurde das Messinstrument stetig überarbeitet. Im Prozess wurde immer wieder Rücksprache mit Rabea Wienholt gehalten.

3.4 Stichprobe

Zunächst kann Bezug auf die Stichprobenart genommen werden, welche bestimmt nach welchem Auswahlverfahren diese ausgewählt wurde. Innerhalb des Projekts handelt es sich um eine nicht-probabilistische Stichprobe. Weitergehend wurde die Auswahl dieser bewusst getroffen, da alle Personen eine bestimmte Merkmalskombination zeigen. Alle erhobenen Personen, sind Schüler*innen von einer der am Projekt teilnehmenden Schulen und somit nicht willkürlich ausgewählt (Döring & Bortz, 2016).

„Der Stichprobenumfang ist definiert als Anzahl aller Objekte oder Personen in der Stichprobe“ (Döring & Bortz, 2016, S.294). Insgesamt nahmen 66 Schüler*innen an den Erhebungen teil. Diese Anzahl lässt sich nochmals in den Ü6 Bereich, mit 30 Schüler*innen und den Ü9 Bereich mit 36 Schüler*innen einteilen. Da die teilnehmenden Kinder sich ausschließlich als männlich und weiblich identifiziert haben, folgt eine binäre Geschlechterdifferenzierung. Insgesamt befanden sich in der Ü6 Stichprobe, 21 männliche und 9 weibliche Personen. Der Ü9 Bereich differenzierte sich in 34 männliche und 2 weibliche Personen. Aufgrund dieser ungleichen Verteilung lassen sich keine geschlechterspezifischen Unterschiede feststellen. Bei den Schüler*innen aus dem Ü6 Bereich wurden insgesamt 30 Testungen innerhalb eines Erhebungszeitpunkts durchgeführt. In dem Ü9 Bereich fanden 67 Testungen, innerhalb von drei Erhebungszeitpunkten statt.

Ebenso muss die Stichprobe im Hinblick auf die Interventionsarten unterschieden werden. Die Interventionsarten setzen sich aus der Bewegungsklasse, Leseklasse, sowie Kontrollklasse zusammen. In der Bewegungsklasse im Ü6 Bereich nahmen 11 Kinder und im Ü9 Bereich 10 Kinder an den Erhebungen teil. In der Leseklasse waren es im Ü6 Bereich 10 und im Ü9 Bereich 14 Kinder. Die Kontrollklasse nahm mit 9 Kindern im Ü6 Bereich und mit 12 Kindern im Ü9 Bereich an dem Projekt teil.

3.5 Untersuchungsdesign

3.5.1 *Qualitatives Forschungsdesign*

In diesem Teil wird auf das qualitative Forschungsdesign eingegangen. Hierfür waren Leitfadeninterviews geplant, die innerhalb der Datenerhebungen in den Schulen mit den Lehrer*innen durchgeführt werden sollten. Ein Leitfadeninterview basiert auf einem vorab systemisch vereinbarten Leitfaden.

„Durch den Leitfaden und die darin angesprochenen Themen erhält man ein Gerüst für Datenerhebung und Datenanalyse, das Ergebnisse unterschiedlicher Interviews vergleichbar macht. Dennoch lässt es genügend Spielraum, spontan aus der Interviewsituation heraus neue Fragen und Themen einzubeziehen oder bei der Interviewauswertung auch Themen herauszufiltern, die bei der Leitfadenskonzepktion nicht antizipiert wurden.“ (Bortz & Döring, 2006, S. 314).

Diese Interviewform war zeitlich nicht möglich, da die Lehrer*innen zu sehr in das Klassengeschehen eingebunden waren und sich somit keine zusätzlichen Zeitslots für die Interviews einräumen konnten. Aufgrund dessen wurde das Forschungsdesign in eine schriftliche Befragung (Anhang B) umgewandelt, um den Vorgang für die Lehrer*innen sowie für die Forschenden realisierbar zu gestalten. Das größte Hauptaugenmerk dieser Befragungsform besteht darin, dass Fragen sowohl schriftlich vorgelegt als auch schriftlich beantwortet werden sollen. Durch das Kriterium des rein schriftlichen Austauschs lassen sich die beiden Formen voneinander abgrenzen. Eine schriftliche Befragung wird in der Wissenschaft auch Fragebogenmethode genannt (Lüdders, 2016). „Unter der wissenschaftlichen Fragebogenmethode („questionnaire method“) verstehen wir die zielgerichtete, systematische und regelgeleitete Generierung und Erfassung von verbalen und numerischen Selbstauskünften von Befragungspersonen zu ausgewählten Aspekten ihres Erlebens und Verhaltens in schriftlicher Form“ (Döring & Bortz, 2016, S.398).

Im Vorhinein wurden drei Kategorien festgelegt, welche jeweils zwei bis drei Fragen beinhalten. Zu den Kategorien zählte zum einen der Effekt (1), die Reaktion und Umsetzung (2) sowie das Verhalten der Kinder (3). Die Kategorie Effekt beinhaltet die

Fragestellungen „Welchen Effekt haben Sie durch die jeweiligen Interventionen gespürt?“ und „Hat sich etwas im Klassengefüge verändert? Wenn Ja, was hat sich verändert?“. Bei der zweiten Kategorie wurden die Fragen „Welche Intervention wurde durchgeführt und in welchem Maße?“, „Gab es ein Feedback von den Eltern? Wenn Ja, welches Feedback?“ und „Würden Sie an der Studie noch mal teilnehmen wollen? Welche Gründe spielen bei der Entscheidung eine Rolle?“ gestellt. Die Fragen: „Wie wurden die Interventionen von den Kindern angenommen? (Hat den Kindern es Spaß gemacht oder waren sie abgeneigt?) und „Welche Veränderung haben die Eltern an Ihren Kindern bemerkt?“ wurden innerhalb der dritten Kategorie gestellt. Hierbei handelt es sich ausschließlich um offene Fragestellungen um eine möglichst große Vielfalt von Informationen zu erlangen. Die Nutzung von offenen Fragen bei einer schriftlichen Befragung ist eher unüblich, jedoch wurden diese genutzt, da es sich zunächst um ein Leitfadeninterview handeln sollte (Züll & Menold, 2014). Aufgrund dessen, dass die zeitliche Verfügbarkeit der Lehrer*innen erst bei den Erhebungen selbst ersichtlich wurde musste das Leitfadeninterview in eine schriftliche Befragung umstrukturiert/angepasst werden. Folglich war zunächst geplant eine Inhaltsanalyse nach Mayring durchzuführen, jedoch wurde dieses Vorgehen verworfen, weil eine zu geringe Datenmenge erhoben wurde. Durch die aktuellen Covid 19 Situation konnten keine weiteren Schulen rekrutiert sowie weitere Erhebungen durchgeführt werden und somit keine schriftliche Befragung gemacht werden. Folglich liegen ausschließlich drei schriftliche Befragungen vor. Demzufolge dienten sie als zusätzliche Informationsquelle zu dem quantitativen Forschungsdesign und wurden nicht explizit ausgewertet, sondern genutzt, um das Projekt reflektieren und ein Ausblick geben zu können. Dementsprechend wurde der angestrebte Mixed-Methods-Ansatz verworfen, da der qualitative Auswertungsprozess nicht stattfinden konnte. Der Mixed-Methods-Ansatz versucht die positiven Seiten der quantitativen- sowie qualitativen Forschung zu kombinieren, um dementsprechend die Forschungsfrage ausführlicher beantworten zu können (Roch, 2017).

3.5.2 *Quantitatives Forschungsdesign*

Das adaptierte Beobachtungsinstrument basiert auf dem fundierten MotorikPlus (Zimmer, 2021), weswegen dieser demnach zur Orientierung genutzt wurde. Im Prozess der Erprobung des MotorikPlus wurde das Verfahren auf Basis der klassischen Testtheorie (KTT) wissenschaftlich untersucht (Zimmer, 2021). Die KTT gehört zu den am häufigsten verwendeten Testtheorien und ist im Wesentlichen als eine Messfehlertheorie zu verstehen (Steyer und Eid, 2001, zitiert nach Moosbrugger, Gäde, Schermelleh-Engel & Rauch, 2020). Mit dieser Feststellung wird bereits deutlich, dass Messungen in der Regel mit einem Fehler behaftet sind. Damit Messungen jedoch in der Psychologie sinnvoll genutzt werden können, beispielsweise für diagnostische Zwecke, sollten sie möglichst genau und zuverlässig sein. Die KTT ist insbesondere für Testitems mit kontinuierlichem, beziehungsweise zumindest vielstufigem Antwortformat konzipiert und konzentriert sich dabei auf die Gewinnung von wahren Merkmalsausprägungen, sogenannten True-Scores, sowie auf die Reliabilität und Validität der Testwerte (Moosbrugger, Schermelleh-Engel, Gäde & Kelava, 2020). Die Reliabilitätsbestimmung ist als wesentliches Ziel der KTT zu verstehen (Moosbrugger, Gäde, Schermelleh-Engel & Rauch, 2020) und ist im Hinblick auf die Beantwortung der Forschungsfrage (Kapitel 3.1) zielführend, weswegen sich insbesondere auf die Errechnung der Reliabilität fokussiert wurde.

Den ersten Schritt der Auswertung stellt die Datenaufbereitung dar. Diese dient dazu das Rohdatenmaterial einer empirischen Studie, welches oftmals in verschiedenen Formen und Formaten existiert, Lücken und Inkonsistenzen enthält und in seiner Gesamtheit schwer zu überblicken ist, aufzubereiten. Zunächst muss sorgfältig sortiert, kommentiert, formatiert, anonymisiert, bereinigt und teilweise auch transformiert werden, um eine systematische Datenanalyse zu ermöglichen. Am Ende der Datenaufbereitung entstehen mit Metainformationen und Kommentierungen versehene, fehlerbereinigte und anonymisierte elektronische Datensätze, welche anschließend mit entsprechendem Auswertungsprogramm auswertbar sind (Döring & Bortz, 2016). Döring und Bortz (2016) nennen dazu verschiedene Schritte, welche im Verlauf der Datenbereinigung berücksichtigt wurden. In einem ersten Schritt wurden somit strukturierte Datensätze für die verschiedenen Altersspannen und Erhebungszeitpunkte

erstellt, welche die Sortierung, Zuordnung, Digitalisierung und Formatierung des Rohdatenmaterials ermöglichen. Es entstanden ein Datensatz für die Altersspanne Ü6 und drei Datensätze für die Altersspanne Ü9. Zusätzlich wurden je Datensatz Subdatensätze für die entsprechenden Kompetenzbereiche erstellt. Der nächste Schritt war die Kommentierung der Datensätze. Darunter ist die Ergänzung der Datensätze um Metainformationen, wie beispielsweise das Erhebungsdatum, zu verstehen. Außerdem werden in quantitativen Datensätzen die Bedeutungen der Variablen und der vergebenen numerischen Werte festgehalten, so zum Beispiel die Codierung des Geschlechts. Als dritter Schritt fand bei der Datenbereinigung die Anonymisierung der Datensätze statt. Hierbei wurden Informationen, die zur Identifizierung von Untersuchungsteilnehmenden führen könnten, entfernt. Innerhalb des Projektes wurden Kürzel zur Anonymisierung erstellt, welche innerhalb des Auswertungsprozesses genutzt wurden. Der vierte Schritt war anschließend die Datenbereinigung. Dabei wurden unplausible und fehlerhafte numerische Angaben identifiziert, so beispielsweise Werte außerhalb des zulässigen Wertebereichs. Außerdem wurde eine sinnvolle Behandlung von Lücken im Datenmaterial durchgeführt, sodass beispielsweise fehlende Werte gekennzeichnet wurden (Döring & Bortz, 2016).

Diese aufbereiteten Datensätze konnten anschließend zur Auswertung mit dem Auswertungsprogramm RStudio genutzt werden. Das Auswertungsprogramm RStudio wurde verwendet, da sich bereits seit längerer Zeit mit dem Programm auseinandergesetzt wurde und somit mehr Kenntnisse mit RStudio, als beispielsweise mit dem Programm SPSS vorlagen. Weiter bietet RStudio flexiblere Analyseschritte und eine individuell anpassbare Ausgabe. Nennenswert ist zusätzlich, dass RStudio für alle gängigen Betriebssysteme anwendbar ist. Um eine adäquate Auswertung mit diesem Programm sicherstellen zu können, ist eine Methodenkenntnis erforderlich, welche sich durch vorbereitende Veranstaltungen und Eigenrecherche angeeignet wurde (Loy, 2015).

Zu Beginn des Auswertungsprozesses hat ein Treffen mit Steffen Wild stattgefunden. Steffen Wild ist Mitarbeiter der TU Dortmund und arbeitet im Fachbereich der Psychologischen Diagnostik. Er besitzt sehr gute statistische Kenntnisse und zeigte erste Möglichkeiten zur Datenauswertung in RStudio auf. Hierbei zeigten sich zunächst Bedienprobleme innerhalb des Programms, welche jedoch überwunden werden

konnten. Zusätzlich wurden die Beratungs- und Analyseleistungen des Bereichs Statistische Beratung und Analyse (SBAZ) des Zentrums für Hochschulbildung der TU Dortmund in Anspruch genommen. Insbesondere im Hinblick auf die Befehlssprache und das damit einhergehende Ablesen der Ausgabe.

Mittels des Programms RStudio wurde somit in Anlehnung an den MotorikPlus (Zimmer, 2021) und die KTT die Reliabilität der Adaption berechnet, indem Cronbachs Alpha, die Itemschwierigkeit und die Trennschärfe ermittelt wurden. Die Ergebnisse der Berechnungen sind in Kapitel 4.3 zu finden. Außerdem wurden Mittelwerte und Standardabweichungen errechnet, um Veränderungen der Leistungen der Schüler*innen zu ermitteln und eventuelle mangelhafte Items, beziehungsweise fehlerhafte Skalierungen aufzudecken. Das Ziel der quantitativen Auswertung ist es eine Aussage über die Güte des Beobachtungsinstruments treffen zu können, insbesondere über die Reliabilität und die damit einhergehende mögliche Verbesserung des Instruments.

3.5.3 Ablauf der Datenerhebungen

In diesem Unterkapitel wird ein Einblick in den allgemeinen Ablauf einer wissenschaftlichen Erhebung gegeben. An dem Tag der Erhebung trafen sich alle pünktlich an der vereinbarten Schule. Die Materialien wurden eigenständig von den Studierenden mitgenommen, sodass diese nach der Erhebung auch bei diesen verblieben. Aufgrund dessen, dass einige mit der Bahn kamen, wurden die Materialien von den Studentinnen mitgenommen, die mit dem Auto zu den Schulen fahren. Als nächstes wurden alle Stationen in einem angemessenen Setting aufgebaut. Damit alle Aufgaben durchgeführt werden konnten, musste Stühle und Tische in den Räumen umgestellt werden. Zu den Stationen wurde die dazugehörige Stationsnummer direkt dazugelegt, damit in der Untersuchung die erhobenen Aufgaben leichter in den Bögen (Anhang D) zu finden waren. Sobald alles aufgebaut war, wurden vereinzelt Schüler*innen aus den Klassen abgeholt. Bei Schulen, bei denen mehrere Klassen erhoben wurden, kamen diese nacheinander dran, sodass immer nur eine Klassenart getestet wurde. Bezüglich der reduzierten Gruppengröße der Projektgruppe, welche sich auf fünf Studentinnen begrenzte, wurden immer nur zwei bis drei Kinder aus der Klasse geholt, damit jedes Kind von einer Studentin getestet werden konnte. Die Dauer einer

einzelnen Testung lag bei 25 bis 30 Minuten. Daraus lässt sich erschließen, dass die Dauer einer Aufgabe bei ein bis zwei Minuten und somit die Gesamtdauer für die wissenschaftliche Erhebung einer Klasse bei circa einer bis anderthalb Stunden liegt. Je nach Auslastung der Klasse erhob jede Studentin innerhalb einer Erhebung von einer Klasse um die zwei bis vier Schüler*innen. Bei der Erhebung hatte jede Studentin einen Durchführungsbogen (Anhang D) dabei, um die genaue Durchführung sowie Bewertungsskala der einzelnen Aufgaben vor Augen zu haben. Gleichzeitig hatte jede Studentin einen Auswertungsbogen, um den Namen, das Geburtsdatum und weitere Informationen des jeweiligen Kindes sowie die entsprechenden Werte dieser in den Bogen eintragen zu können. Nachdem die Kinder getestet wurden, liefen sie wieder eigenständig in ihre Klasse zurück. Bevor die nächsten Kinder aus der Klasse geholt wurden, machten sich die Studentinnen Notizen, bei denen die Werte in den Auswertungsbogen nachgetragen oder auffälligen Verhaltensweisen und Äußerungen am Rand notiert wurden. Sobald alle Kinder getestet wurden, bauten die Studentinnen die Stationen ab, der Raum wurde durchgefegt und die Tische sowie Stühle wurden wieder zurückgestellt. Nach der wissenschaftlichen Erhebung trug jede Studentin die Werte in die dafür vorgesehene Rohwerttabelle, für die getesteten Kinder, ein. Zudem erfolgte eine Nachbesprechung der Erhebung, um diese reflektieren zu können und Anmerkungen in den Durchführungsbögen zu vermerken, wodurch diese verbessert werden konnten.

4. Darstellung der quantitativen Forschungsergebnisse

Um die Qualität wissenschaftlicher Forschung beurteilen zu können, gibt es Gütekriterien, welche „[...] [die] Mindestanforderungen [definieren], denen ein empirisches Forschungsprojekt zu genügen hat, wenn es als wirklich gute wissenschaftliche Forschung gelten möchte“ (Hug & Poscheschnik, 2020, S.115). Sowohl bei quantitativen als auch bei qualitativen Test- sowie Beobachtungsverfahren ist die Sicherung der Güte der Instrumente zur Datengewinnung und der Resultate Gegenstand methodologischer Reflexionen (Pfeiffer & Püttmann, 2018). Wissenschaftliche Testverfahren müssen insofern hohe Qualitätsanforderungen entsprechen. Demnach sollte die Entwicklung und Gestaltung der Fragebögen und Testsituationen bei Beobachtungs- sowie Testverfahren möglichst viele der mittels Gütekriterien angeforderten Qualitätsansprüchen genügen. Des Weiteren versteht man „[u]nter dem Begriff „Gütekriterien“ [...] eine Reihe von Gesichtspunkten/Anforderungen, die bei der Test- und Fragebogenkonstruktion zur Qualitätssicherung Berücksichtigung finden sollen“ (Moosbrugger & Kelava, 2020, S.17). Die als Hauptgütekriterien bekannten Objektivität, Reliabilität und Validität und Nebengütekriterien liefern nicht nur eine Beurteilung über die Qualität eines diagnostischen Verfahrens, sondern unter anderem auch Anhaltspunkte für den Wahrheitsgehalt sowie die Haltbarkeit der wissenschaftlichen Aussagen.

4.1 Objektivität

Um die notwendige Vergleichbarkeit der Ergebnisse verschiedener Proband*innen in Tests und Fragebögen zu gewährleisten, muss das Qualitätskriterium der Objektivität erfüllt sein. Demensprechend bedeutet, dies, „[...] dass dem Testleiter kein Verhaltensspielraum bei der Durchführung, Auswertung und Interpretation des Tests bzw. Fragebogens eingeräumt wird“ (Moosbrugger & Kelava 2020, S.18). Hug und Poscheschnik (2020) zufolge bedeutet Objektivität in der empirischen Forschung das Bemühen, eine Studie so durchzuführen, dass die Ergebnisse möglichst wenig durch Vorurteile der Forscher*innen verzerrt werden.

Hier wird deutlich, dass Testverfahren erst dann objektiv sind, wenn die Messergebnisse unabhängig von der messenden Person und der Messsituationen sind. Moosbrugger und Kelava (2020) definieren Objektivität wie folgt:

“Ein Test ist dann objektiv, wenn das ganze Verfahren, bestehend aus Testmaterialien, Testdarbietung, Testauswertung und Interpretationsregeln, so genau festgelegt ist, dass der Test unabhängig von Ort, Zeit, Testleiter und Auswerter durchgeführt werden könnte und für eine bestimmte Testperson bezüglich des untersuchten Merkmals dennoch dasselbe Ergebnis und dieselbe Ergebnisinterpretation liefert“ (S.18).

Demgemäß ist beispielsweise eine Beobachtung in der ersten Erhebungsphase sozialer und emotionaler Fähigkeiten während dem Bruderhilfe Spiel erst dann objektiv, wenn verschiedene Beobachter*innen zum gleichen Ergebnis hinsichtlich der Anzahl und Deutung der sozialen Handlungen kommen. Im Zusammenhang zur unserer Testdurchführung, ist die Erfüllung der Durchführungsobjektivität verbesserungswürdig, da aufgrund der Gruppengröße immer nur ein*e Beobachter*in für eine Testperson bestimmt wurde. Andererseits ist zu benennen, dass ein standardisiertes Erhebungsverfahren entwickelt und eingesetzt wurde.

4.1.1 Durchführungsobjektivität

Die Durchführungsobjektivität stellt sicher, dass unabhängig davon, wer die Messung vorgenommen hat die gleichen Messergebnisse erscheinen. Erreicht wird dies erst dann, wenn eine hohe Standardisierung der Messinstrumente, also der Frage- und Testbögen, stattfindet. Moosbrugger und Kelava (2020) definieren, dass die Durchführungsobjektivität vorliegt, wenn die Durchführungsbedingungen in der Weise standardisiert sind, dass das Testverhalten der Testperson nur von der individuellen Ausprägung des interessierenden Merkmals abhängt. Darüber hinaus sollen alle anderen Durchführungsbedingungen konstant oder kontrolliert sein, damit sich diese nicht störend und ergebnisverzerrend auf die Testung auswirken können. Um keine verfälschten Messergebnisse zu erhalten, müssen daher sowohl bei der Erstellung der Testmaterialien als auch bei der Bestimmung der Zeitdauer sowie der Instruktion für jede Aufgabe alle Störeinflüsse eliminiert werden. In der Instruktion wird den Testpersonen

ihre Aufgabe, möglichst schriftlich erklärt; hierbei hat sich die Bearbeitung einiger gleichartiger Probeitems als sehr hilfreich erwiesen. Es muss auch festgelegt werden, ob und wie etwaige Fragen der Testpersonen zum Test behandelt werden sollen. Üblicherweise werden Fragen durch Rückverweis auf die Instruktion beantwortet, weshalb dort alles Wesentliche enthalten sein sollte. In den ersten Testsituationen wurden diese Anforderung nicht erfüllt, da einige Beobachter*innen, die Anleitungen wiederholten, jedoch andere die Aussage: „Mach es so wie du es verstanden hast“ verwendeten, welches zu einer Verzerrung der Testsituation führte. Nach Moosbrugger und Kelava (2020) ist für die Förderung der Durchführungsobjektivität signifikant auf eine Interaktion zwischen Testleiter*in und Testperson zu verzichten oder diese zu minimieren, um Testleiter*inneneinflüsse (z. B. unterschiedliche Betonungen) zu vermeiden. Es wurde darauf geachtet die Instruktionen sehr präzise und möglichst genau für jede Beobachtungsperson zu formulieren sowie bei jeder Testperson gleichermaßen die Anleitung einzuführen. Jedoch scheint die Betrachtung der Durchführungsobjektivität an dieser Stelle für die Durchführung des Testmanuals als verbesserungswürdig, da beispielsweise die Anleitung der Aufgabe 9 im Durchführungsbogen wie folgt lautet: „Hier siehst du vier Behälter, die alle unterschiedlich schwer sind. Sortiere sie der Schwere nach. Bei der Feder liegt das Leichteste und bei dem Elefanten das Schwerste.“ Somit wird deutlich, dass zwei verschiedene Aspekte, die von der Testleitung betont werden „Sortiere sie der Schwere *nach...*“ oder „Bei der Feder liegt das **Leichteste** und bei dem Elefanten das **Schwerste**“ zwei diverse und somit verfälschte Verhaltensergebnisse hervorrufen. Dieser Aspekt zeigt, dass in einer schriftlichen computerbasierten Instruktion mit hoher Wahrscheinlichkeit eine Interaktion hätte vermieden werden können und somit eine eigene Betonung der Testleitung das Verhalten der Testpersonen beeinflusst. Nennenswert erscheint hier, dass jede Testperson bei jeder Station dieselben Materialien sowie Bedingungen zur Verfügung hatte, welche vor jeder Erhebungssituation vorbereitet und kontrolliert wurden. Das bedeutet, dass bei Station 3. „Wurf auf Zielscheibe“ bei jeder Erhebung, die Entfernung von der Wurfscheibe bis zur Markierung abgemessen und bei jeder Testperson identisch durchgeführt wurde. Außerdem wurde bei der Station 7. „rechts-links Liniensprung“ für die Messung der Ausdauer jeder Testperson, die Zeit gestoppt und die Testfläche, mit den Maßen 50x100 cm einschließlich der Mittellinie, bereits

vor der Durchführung vorbereitet. Bei der Station 14. „Vierecke Formen“ bekamen die Testpersonen für die Testung der Feinmotorischen Fähigkeiten auch dieselbe Anzahl an Knetbällen und Zahnstochern, um daraus ein Viereck zu formen.

4.1.2 Auswertungsobjektivität

Die Auswertungsobjektivität eines Tests ist dann gegeben, wenn es eine eindeutige Anweisung gibt, wie die Antworten der Testperson auf die einzelnen Testaufgaben hinsichtlich der Unterscheidung von hohen beziehungsweise niedrigen Merkmalsausprägungen zu kodieren sind (Moosbrugger & Kelava, 2020). Zudem erfordert die Auswertungsobjektivität, „[...] dass eine Methode so konstruiert ist, dass unterschiedliche Forscher zum selben Ergebnis kommen“ (Hug & Poscheschnik, 2020, S.116). Daher wurden für das Beobachtungsinstrument nach der ersten Erhebung genaue Anleitungen sowie Skalenentwürfe im Manual erstellt. Um die Auswertungsobjektivität zu gewährleisten wurden außerdem möglichst verschiedene Merkmalsausprägungen und Verhaltensweisen im Auswertungsbogen ausgeführt, die den Itemwerten zugeordnet wurden. Die Itemwerte wurden während der Durchführung ohne weitere Erschwernis von den Testleitern*innen notiert. Entscheidend ist hier, dass bei der Auswertung der Testergebnisse unterschiedliche Auswerter*innen zu denselben Ergebnissen kommen müssen. Hierbei ist zu erwähnen, dass quantitative Daten keine großen Schwierigkeiten bei der Vergabe von Itemwerten. Bei der Auswertung wurde bemerkt, dass die Vergabe der Itemwerte bei numerischen Daten, wie die Anzahl von Sprüngen oder Zeitspannen deutlich objektiver war als beispielsweise bei der Vergabe von Itemwerten für die Beobachtung der Frustrationstoleranz.

4.1.3 Interpretationsobjektivität

“Die Interpretationsobjektivität fordert schließlich, dass die mithilfe der Methode gewonnenen Werte von unterschiedlichen Forscher*innen gleich interpretiert werden“ (Hug und Poscheschnik, 2020, S.117). Die Interpretationsobjektivität liegt vor, wenn verschiedene Testanwender*innen gleiche Testwerte von verschiedenen Testpersonen bezüglich des untersuchten Merkmals in gleicher Weise interpretieren. Im weiteren Verlauf ist eine quantitative und qualitative Auswertung vorgesehen, um die Entwicklung der Kinder dokumentieren und interpretieren zu können.

4.2 Validität

„Beim Gütekriterium der Validität [...] handelt es sich hinsichtlich der praktischen Anwendung von Tests und der theoretischen Diskussion von Merk- zusammenhängen um das wichtigste Gütekriterium überhaupt, wobei eine hohe Objektivität und eine hohe Reliabilität zwar notwendige, aber keine hinreichen- den Bedingungen für eine hohe Validität darstellen“ (Moosbrugger & Kelava, 2020, S.30). Die Validität eines Tests bezeichnet das Ausmaß, in dem das Messinstrument tatsächlich das zu mes- sende Merkmal (und nicht irgendein anderes) und somit den Grad der Gültigkeit wis- senschaftlicher Aussagen erfasst (Pfeiffer & Püttmann, 2018, Döring & Bortz, 2016). Bei der Validität wird zwischen vier verschiedenen Typen unterschieden: Konstrukt- validität, interne – und externe Validität sowie statistische Validität. Für unser Erhe- bungsverfahren wurde das Testmanual Motorik Plus von Zimmer (2021) verwendet und adaptiert, welches hinsichtlich der Inhalte, Struktur und Praxistauglichkeit als va- lide gilt und daher der inhaltlichen Validität entspricht (Zimmer, 2021). Somit wurden Aufgaben aus validen Tests, wie dem MotorikPlus sowie anderen Tests, genutzt, bei vollendeter Überarbeitung des Erhebungsinstrument sollte eine erneute Validitätsprü- fung erfolgen.

4.3 Reliabilität

Die Reliabilität ist ein Maß dafür, wie genau ein wissenschaftlicher Test das Merkmal, welches beobachtet werden soll, untersucht: „Beobachtungsverfahren sind reliabel, wenn die zu beobachtenden Merkmale – in MotorikPlus die Entwicklungs- und Kom- petenzbereiche – zuverlässig dokumentiert und nicht von situativen Bedingungen (z.B. Tageszeit, Verhalten des Kindes, Beziehung zwischen pädagogischer Fachkraft und Kind) beeinflusst werden“ (Zimmer, 2021, S.51). Um die Reliabilität zu untersu- chen, wird die interne Konsistenz für jeden Kompetenzbereich über den Cronbachs Alpha-Koeffizienten berechnet: „Als Maß für die Güte einer Skala wird ihre Reliabilität (Genauigkeit der Messung) berechnet. Dazu wird der Alpha-Koeffizient nach Cron- bach bestimmt. Alpha ist ein sehr häufig benutztes Maß, das Werte zwischen 0 (un- reliabel) und 1 (perfekte Reliabilität) annehmen kann“ (Kuckartz, 2013, S.247).

Ebenso Zimmer untersuchte die Reliabilität ihres Beobachtungsverfahrens durch die Berechnung der Cronbachs-Alpha-Koeffizienten für jeden Entwicklungs- sowie Kompetenzbereich (Zimmer, 2021). Deshalb wurde dasselbe Verfahren für diese Adaption genutzt. Dazugehörig folgt die Untersuchung der Eignung der jeweiligen Items. Zimmer kontrollierte diese durch die Berechnung der Itemschwierigkeit sowie der Itemtrennschärfe (Zimmer, 2021). Auch nach Kuckartz (2013) sind die Itemschwierigkeit und die Itemtrennschärfe besonders aussagekräftig zur Beurteilung der Eignung von Items und somit auch für die Bewertung der Reliabilität. Daher wurden bei der Untersuchung der Reliabilität für die Adaption des MotorikPlus, ebendiese Koeffizienten berechnet.

„Die Trennschärfe bzw. der Trennschärfekoeffizient gibt an, wie gut ein einzelnes Item das Zielkonstrukt des Tests misst, bzw. wie hoch das Item mit dem Gesamtwert korreliert“ (Döring & Bortz, 2016, S.477). Folgend kann einem trennscharfen Item durch die Beantwortung ebendieses Items herausgelesen werden, wie gut oder schlecht die Testperson im Gesamtest abschneiden wird.

„Die Schwierigkeit eines Items („item difficulty“) wird durch einen Schwierigkeitsindex ausgedrückt. Ein Item ist umso schwieriger, je geringer seine Zustimmungs- bzw. Lösungsrate in der betrachteten Stichprobe ausfällt. Umgekehrt wird ein leichtes Item von fast allen Testpersonen in Schlüsselrichtung beantwortet bzw. korrekt gelöst“ (Döring & Bortz, 2016, S.477). Die Itemschwierigkeit wurde von Zimmer nach Fisseni, 1990, berechnet (Zimmer, 2021). Es wurde in der Berechnung der Itemschwierigkeit dieselbe Methode herangezogen.

4.3.1 Itemschwierigkeiten

„Extrem schwierige Items, denen kaum jemand zustimmt (z. B. nur 0% – 20% Zustimmung/Löser), oder extrem leichte Items, die von fast allen Testpersonen gelöst werden (z. B. 80% – 100% Zustimmung/Löser), sind wenig informativ, da sie keine Personenunterschiede sichtbar machen“ (Döring & Bortz, 2016, S.477). Demnach muss der Wert der Itemschwierigkeit zwischen 0.2 und 0.8 liegen, damit die Items eine angemessene Schwierigkeit haben. Für einen wissenschaftlichen Test oder Beobachtungsverfahren ist es zusätzlich wichtig, dass die Items unterschiedliche Schwierigkeitsgrade haben: „Damit ein Test die verschiedenen Testpersonen mit ihren

unterschiedlichen Fähigkeiten annähernd gleich gut differenziert, ist darauf zu achten, dass die Items eine möglichst breite Schwierigkeitsstreuung aufweisen, dass im Test also leichte bis schwierige Items vertreten sind.“ (Döring & Bortz, 2016). Wie in der allgemeinen Beschreibung bereits erwähnt, wurde sich für die Berechnung der Itemschwierigkeit nach der Definition von Fisseni, 1990, gerichtet: „Der Schwierigkeitsindex p ist definiert für zweistufige Antworten. Sehen Item-Antworten mehrere Abstufungen vor, etwa Werte von 0 bis 10, dann ist p zur Angabe der „Schwierigkeit“ nicht definiert. Um dennoch Angaben zu gewinnen, kann man unterschiedlich vorgehen: [...] Man berechnet einen eigenen Wert für p : Er sei „mehrstufiges p “ genannt, sein Kürzel sei „ pm “ (Fisseni, 1990, S. 34) Um pm zu berechnen wird die Summe aller erreichten Werte aller Testpersonen, durch die Summe der höchstmöglichen Werte aller Testpersonen dividiert (Fisseni, 1990). Folgend wird die Rechnung an einem Beispiel (Item: Wurf auf Zielscheibe, zweite Erhebung Ü9) erläutert. Für die Summe der erreichten Werte aller Schüler*innen wurden alle erreichten Werte der Schüler*innen addiert und ergeben 56. Dann wurde die Summe der höchsten Werte berechnet. Es nahmen 24 Schüler*innen teil und der höchste Wert des Items liegt bei vier. Demnach wurde 24 mit vier multipliziert. Das ergab 92. Im letzten Schritt wurde dann die 56 durch die 92 dividiert. Das Ergebnis lautet 0.61 und das Item ist somit mittelschwer. Nun wird auf die Ergebnisse der Auswertung der Itemschwierigkeit eingegangen.

Kompetenzbereich	Ü6/1	Ü9/1	Ü9/2	Ü9/3
Koordination	0.4-0.86	0.61-0.95	0.61-0.65	0.6-0.64
Kraft, Ausdauer	0.53-0.8	0.57-0.69	0.53-0.89	0.65-0.97
Gleichgewicht	0.48-0.8	0.48	0.71-0.87	0.46-0.85
Feinmotorik	0.53-0.84	0.56-0.85	0.45-0.78	0.74-0.88
Taktil-kinästhetische Wahrnehmung	0.65-0.78	0.82-0.88	0.91	0.96-0.97
Visuelle Wahrnehmung	0.65-0.8	0.79-0.91	0.78-0.95	0.89-1
Emotionale Kompetenzen	0.73	0.66	0.8	0.93
Soziale Kompetenzen	0.75-0.77	0.85	0.74-0.79	0.89-0.9
Kognitive Kompetenzen	0.68-0.7	0.7	0.78-0.86	0.83-0.94

Abb. 1

Anmerkungen. Tabelle: Itemschwierigkeiten.

In der Adaption für die Altersspanne Ü6 sind die Items mittelschwer bis leicht. Die Itemschwierigkeit ist bei den Kompetenzbereichen Koordination (0.4-0.86) und Feinmotorik (0.53-0.84) zu einfach. Zwar ist die Schwere der Items teilweise mittel, jedoch haben beide jeweils ein Item mit einer zu geringen Schwierigkeit. Die Kompetenzbereiche Kraft & Ausdauer (0.53-0.8) und Gleichgewicht (0.48-0.8) weisen beide eine gute Verteilung der Itemschwierigkeit auf, obwohl das leichteste Item mit jeweils 0.8 bereits grenzwertig leicht ist.

In der ersten Adaption für die Altersspanne Ü9 sind die Items auch mittelschwer bis leicht. Ein sehr schweres Item liegt in der Adaption nicht vor. Teilweise zu einfach sind die Items der Kompetenzbereiche Koordination (0.61-0.95), Feinmotorik (0.56-0.85), taktil-kinästhetische Wahrnehmung (0.82-0.88), visuelle Wahrnehmung (0.79-0.91), soziale Kompetenzen (0.85) und auditive Kompetenzen (0.82). Die Kompetenzbereiche Kraft & Ausdauer (0.57-0.69), Gleichgewicht (0.48) und emotionale Kompetenzen (0.66) weisen angebracht schwierige Items auf. Dazu kommt, dass eine hohe Schwierigkeitsstreuung vorliegt. Das ist positiv für die Differenzierung der Testpersonen (Döring & Bortz, 2016).

Die Items der Kompetenzbereiche der zweiten Adaption für die Altersspanne Ü9 sind sowohl schwieriger, als auch leichter geworden. Schwerer geworden sind die Items der Kompetenzbereiche Koordination (0.61-0.95 -> 0.61-0.65) Kraft & Ausdauer (0.57-0.69 -> 0.53-0.89), Feinmotorik (0.56-0.85 -> 0.45-0.78) und soziale Kompetenzen (0.85 -> 0.74-0.79). Leichter geworden sind die Items der Kompetenzbereiche Gleichgewicht (0.48 -> 0.71-0.87), taktil-kinästhetische Wahrnehmung (0.82-0.88 -> 0.91), emotionale Kompetenzen (0.66 -> 0.8) und kognitive Kompetenzen (0.7 -> 0.78-0.86). Die Items des Kompetenzbereichs taktil kinästhetische Wahrnehmung sind zu leicht, denn die Itemschwierigkeit liegt bei 0.91. Die Kompetenzbereiche Kraft & Ausdauer, Feinmotorik und kognitive Kompetenzen haben jeweils mittelschwere bis leichte Items, aber auch jeweils ein Item, welches zu leicht ausfällt (0.53-0.89; 0.71-0.89; 0.78-0.86). Die Items der Kompetenzbereiche Koordination (0.61-0.65) und Feinmotorik (0.45-0.78) haben beispielsweise gute Itemschwierigkeiten.

Durch die dritte Erhebung für die Altersspanne Ü9 haben sich die Itemschwierigkeiten noch einmal verändert. Die Schüler*innen hatten mehr Schwierigkeiten bei den Items

der Kompetenzbereiche Gleichgewicht (0.71-0.87 -> 0.46-0.85). Alle anderen Items konnten von den Schüler*innen besser bearbeitet werden, wodurch die Itemschwierigkeit bei allen anderen Kompetenzbereichen sinkt. Dies hat zur Folge, dass fast alle Items zu einfach sind. Lediglich Items der Kompetenzbereiche Koordination (0.6-0.64), Kraft & Ausdauer (0.65-0.97), sowie Gleichgewicht (0.46-0.85) sind teilweise schwer genug. Jedoch weisen auch hier nicht alle Items eine angebrachte Schwierigkeit auf, lediglich der Kompetenzbereich Koordination. Die Tatsache, dass die Items fast alle besser „beantwortet“ werden konnten, kann daran liegen, dass die Items nicht mehr verändert wurden nach der zweiten Erhebung und sich die Schüler*innen erinnern konnten. Besonders im kognitiven Bereich konnte auch schon während der Erhebung festgestellt werden, dass sich die Schüler*innen an die Antworten erinnern, was zu einer höheren Lösungsquote und somit zu einer geringeren Itemschwierigkeit führen kann.

Im Gesamten sind die Items demnach zu einfach. Entweder müsste die Bewertung nochmals angepasst, oder überarbeitete und neu erstellte Items genutzt werden.

4.3.2 Cronbachs Alpha

Am häufigsten wird zur Berechnung der Reliabilität von Tests die Ermittlung der internen Konsistenz genutzt. Mithilfe der internen Konsistenz wird der Anteil der gemeinsamen Varianz an der Gesamtvarianz der jeweils relevanten Items bezeichnet (Krumm, Schmidt-Atzert & Amelang, 2021). Auch Zimmer (2021) berechnete zur Überprüfung der Reliabilität der einzelnen Entwicklungs- und Kompetenzbereiche des MotorikPlus die interne Konsistenz mithilfe von Cronbachs Alpha. Döring und Bortz (2016) geben an, dass es sinnvoll ist, die interne Konsistenz über einen Cronbachs Alpha-Koeffizienten für jede Subskala einzeln zu bestimmen, statt für alle Items gemeinsam einen Alpha-Koeffizienten zu berechnen. Folglich wurde der Alpha-Koeffizient für jeden Kompetenzbereich individuell berechnet. Cronbachs Alpha kann Werte zwischen 0 und 1 annehmen. Der Alpha-Koeffizient ist umso höher, desto mehr Items ein Test oder Untertest enthält und je höher die Item-Interkorrelationen ausfallen (Döring & Bortz, 2016). Als Orientierung zur Interpretation von Cronbachs Alpha nennt Blanz (2021) folgende Angaben. Werte $>.9$ sind exzellent, Werte $>.8$ können als gut bezeichnet werden, Werte $>.7$ sind akzeptabel, Werte $>.6$ fragwürdig, Werte

>.5 sind als schlecht zu beurteilen und Werte <.5 als inakzeptabel. Kuckartz, Rädiker, Ebert & Schehl (2013) geben an, dass in bestimmten Kontexten Skalen mit Alpha-werten zwischen 0.7 und 0.8 als brauchbar, Werte ab 0.8 als gut und Werte ab 0.9 als sehr gut einzustufen sind. Wichtig zu nennen ist zudem, dass Cronbachs Alpha, wie alle Reliabilitätskoeffizienten, stichprobenabhängig ist (Krumm, Schmidt-Atzert & Amelang, 2021). Daher sind die Werte nicht direkt miteinander zu vergleichen, da die Anzahl der Proband*innen zu den verschiedenen Erhebungszeitpunkten schwankt.

Im Folgenden wird auf die Ergebnisse der Auswertung von Cronbachs Alpha eingegangen, indem diese eigeordnet werden.

Kompetenzbereich	Ü6/1	Ü9/1	Ü9/2	Ü9/3
Koordination	0.39	0.29	0.39	0.45
Kraft, Ausdauer	0.38	0.73	0.39	0.57
Gleichgewicht	0.49	0.63	0.66	0.45
Feinmotorik	0.18	-	0.23	0.4
taktil-kinästhetische W	0.53	0.1	0.04	0.69
visuelle Wahrnehmung	0.12	0.41	0.05	-
emotionale Kompetenzen	-	-	-	-
soziale Kompetenzen	0	-	0.08	0.5
kognitive Kompetenzen	0.03	0.47	0.37	0.1
auditive Kompetenzen	-	-	-	-
Median	0.28	0.44	0.3	0.45

Abb. 2

Anmerkungen. Tabelle: Cronbachs Alpha.

Der Cronbachs- α -Koeffizient fällt für die Adaption Ü6 (mit einer Ausnahme) inakzeptabel aus (Ü6/1: Median Cronbachs α = .28, α -Range 0 - 0.53). Hierbei ist zu beachten, dass für diese Altersspanne nur die erste Erhebung stattfinden konnte, weswegen keine Daten der Überarbeitung vorhanden sind. Für die Altersspanne Ü9 fällt der Cronbachs- α -Koeffizient innerhalb der ersten Erhebung mit zwei Ausnahmen inakzeptabel aus. Die Ausnahmen befinden sich zum einen im akzeptablen und zum anderen im fragwürdigen Bereich (Ü9/1: Median Cronbachs α = .44, α -Range 0.1 – 0.73). Innerhalb des zweiten Erhebungszeitpunktes zeigen sich ein Alpha Wert im Bereich des inakzeptablen, mit Ausnahme eines Wertes, welcher sich im fragwürdigen Bereich befindet (Ü9/2: Median Cronbachs α = .3, α -Range 0.04 - 0.66). Der Cronbachs- α -Koeffizient des dritten Erhebungszeitpunktes fällt mit drei Ausnahmen inakzeptabel aus, zwei Werte sind als schlecht zu beurteilen und ein Wert als fragwürdig (Ü9/3:

Median Cronbachs $\alpha = .45$, α -Range 0.1 - 0.69). Insgesamt fallen die Werte eher schlecht aus, was unter anderem an der geringen Itemanzahl liegen könnte. Da es sich um ein alltagstaugliches Beobachtungsinstrument handelt, wurde versucht die Itemanzahl möglichst gering zu halten, um eine einfachere Durchführbarkeit zu gewährleisten. Dies bedingt zum Teil die geringen Alpha-Koeffizienten, da diese umso höher ausfallen, desto mehr Items ein Instrument enthält.

In der Tabelle fällt auf, dass nicht für alle Kompetenzbereiche die interne Konsistenz ermittelt werden konnte. Dies liegt im Falle der „auditiven Kompetenzen“ und „emotionalen Kompetenzen“ daran, dass die Skalen über zu wenig Items verfügen, da mindestens zwei Items vorhanden sein müssen. Weiter ist zu dem Kompetenzbereich „visuelle Wahrnehmung“ zu nennen, dass die Schüler*innen der Altersspanne Ü9 während der dritten Erhebung in einem Item durchgehend den Wert 4 erzielten, weswegen es keine Varianz innerhalb des Items gab.

4.3.3 Trennschärfe Koeffizienten

Neben der Itemschwierigkeit kann als weiterer Itemkennwert die Trennschärfe für jedes Item angegeben werden. Die Trennschärfe ist die bivariate Korrelation des Items mit dem Gesamtscore der Skala (Döring & Bortz, 2016). Die Trennschärfe zeigt an, in welchem Ausmaß das Item das Gleiche misst wie der Test, beziehungsweise bei mehrdimensionalen Tests die Subskala des Tests (Krumm, Schmidt-Atzert & Amelang, 2021). Sie ist ein Indikator dafür, wie gut das betreffende Item die Skala repräsentiert (Kuckartz, Rädiker, Ebert & Schehl, 2013). Eine hohe Trennschärfe bedeutet folglich, dass das Item gut zwischen Personen mit hohem und solchen mit niedrigem Skalenwert differenzieren kann (Döring & Bortz, 2006, zitiert nach Blanz, 2021). Es kann zwischen der unkorrigierten und der korrigierten Trennschärfe unterschieden werden. Insbesondere dann, wenn der Test nur aus wenigen Items besteht, wird die unkorrigierte Itemtrennschärfe dadurch überschätzt, dass der jeweilige Itemwert auch selbst in den Testwert einfließt, was sich korrelationserhöhend auswirkt. Daher empfiehlt sich bei geringer Itemanzahl die Berechnung der korrigierten Trennschärfe. Auch Zimmer berechnete im MotorikPlus (2021) die korrigierte Trennschärfe der Items. Bei der korrigierten Trennschärfe wird die sogenannte „part-whole-correction“ durchgeführt, bei der das betreffende Item bei der Summenbildung ausgelassen wird

(Kelava & Moosbrugger, 2020). Die korrigierte, als auch unkorrigierte Trennschärfe kann Werte von -1 bis $+1$ annehmen, dabei gilt je höher die Werte im positiven Bereich liegen, desto besser. Positive Werte zwischen $.30$ und $.50$ können als mittelmäßig und Werte größer als $.50$ als hoch eingestuft werden (Döring & Bortz, 2016). Gängige Konventionen empfehlen, möglichst nur Items mit einer Trennschärfe größer als $.30$ in einem Test zu belassen (Krumm, Schmidt-Atzert & Amelang, 2021). Items mit hohen Trennschärfewerten führen zu einem hohen Cronbachs Alphawert für die Gesamtskala (Kuckartz, Rädiker, Ebert & Schehl, 2013).

Im Folgenden wird nun auf die Ergebnisse der Auswertung der korrigierten Trennschärfe eingegangen, indem diese eingeordnet werden.

Kompetenzbereich	Ü6/1	Ü9/1	Ü9/2	Ü9/3
Koordination	0.01-0.31	0.05-0.27	0.21-0.36	0.3
Kraft, Ausdauer	0.1-0.38	0.52-0.57	0.25	0.22-0.58
Gleichgewicht	0.08-0.56	0.32-0.48	0.46-0.48	0.12-0.41
Feinmotorik	0.09-0.23	/	0.17	0.3
taktil-kinästhetische W	0.36	0.51	0.02	0.59
visuelle Wahrnehmung	0.07	0.38	0.04	/
emotionale Kompetenzen	/	/	/	/
soziale Kompetenzen	0	/	0.05	0.33
kognitive Kompetenzen	0.02	0.32	0.23	0.06
auditive Kompetenzen	/	/	/	/

Abb. 3 Anmerkungen. Tabelle: korrigierte Itemtrennschärfe.

Die untere Grenze der Trennschärfe-Koeffizienten für die Items der Adaption Ü6 liegen zum Teil nahe null ($0.01-0.1$), wodurch deutlich wird, dass diese Items schlecht zwischen Personen mit hohem und niedrigem Skalenwert differenzieren. Die Kompetenzbereiche Koordination, Kraft/Ausdauer, Gleichgewicht und taktil-kinästhetische Wahrnehmung weisen hingegen zum Teil Trennschärfe-Koeffizienten im mittelmäßigen und sogar hohen Bereich auf ($0.36-0.56$), welche darauf hinweisen, dass die betreffenden Items die Skala gut repräsentieren. Die Trennschärfe-Koeffizienten der ersten Erhebung der Altersspanne Ü9 weisen, bis auf den Kompetenzbereich Koordination, nur Werte im mittelmäßigen bis guten Bereich auf ($0.32-0.57$), wodurch deutlich wird, dass eine gute Differenzierung zwischen Personen mit hohem und niedrigem Skalenwert stattgefunden hat. Die Werte der Trennschärfe-Koeffizienten der zweiten Erhebung Ü9 zeigen geringere Werte im Vergleich zu der ersten Erhebung.

Nur der Kompetenzbereich Gleichgewicht weist keine Werte im schlechten Bereich auf. Alle anderen Kompetenzbereiche weisen Koeffizienten unter .30 auf. Nur die Kompetenzbereiche Gleichgewicht und Koordination weisen einen oder mehr Werte im Bereich des mittelmäßigen vor. Kein Wert liegt im guten Bereich. Die Trennschärfe-Koeffizienten der Altersspanne Ü9 der dritten Erhebung zeigen bis auf den Kompetenzbereich kognitive Kompetenzen mittelmäßige Werte, die Kompetenzbereiche Kraft/Ausdauer und taktil-kinästhetische Wahrnehmung weisen Werte im guten Bereich auf (0.58-0.59).

Für die Kompetenzbereiche „emotionale Kompetenzen“ und „auditive Kompetenzen“ konnten keine Trennschärfe-Koeffizienten errechnet werden, da diese Skalen jeweils nur ein Item enthielten. Außerdem ist auch hier die Stichprobe ein entscheidender Faktor, ebenso wie bei der Berechnung von Cronbachs Alpha. Krumm, Schmidt-Atzert, & Amelang (2021) geben als Beispiel an, dass wenn ein Test einer sehr homogenen Stichprobe von Personen vorgegeben wurde, bei der das Merkmal nicht stark variiert, die Trennschärfen zwangsläufig niedriger ausfallen müssen als bei einer heterogenen Stichprobe.

5. Überarbeitung Manual

5.1 Ü6 Bereich

Im Folgenden Abschnitt geht es um die finale Änderung des Durchführungsmanuals anhand der Ergebnisse der quantitativen Analyse mit RStudio. Zunächst wird auf das Manual aus dem Ü6 Bereich eingegangen. Dieses wird im Hinblick auf Cronbachs Alpha, die Itemmittelwerte, -schwierigkeiten und -trennschärfe überarbeitet.

Nach der eigentlichen Differenzierung der Werte von Cronbachs Alpha können keine Items (Aufgaben) aussortiert werden. Die Items wurden hierbei in die einzelnen Kompetenzbereiche eingeteilt. Demnach wären die Werte 0.7 bis 0.8 brauchbar, 0.8 gut und 0.9 sehr gut. Kein Kompetenzbereich entspricht einem brauchbaren Wert. Dies liegt jedoch daran, dass die Werte bei einer geringeren Item Anzahl schwächer ausfallen. Der Kompetenzbereich emotionale Kompetenzen enthält lediglich ein Item, weshalb die Berechnung von Cronbachs Alpha nicht möglich war. Auch die sozialen- und kognitiven Kompetenzen haben Werte nahe 0 und müssten demnach aus dem Erhebungsinstrument gestrichen werden. Hinsichtlich der unbrauchbaren Werte, sollten folgend Items ergänzt werden, statt das komplette Erhebungsinstrument zu verwerfen (Kuckartz et. al. 2013) (Abb. 4).

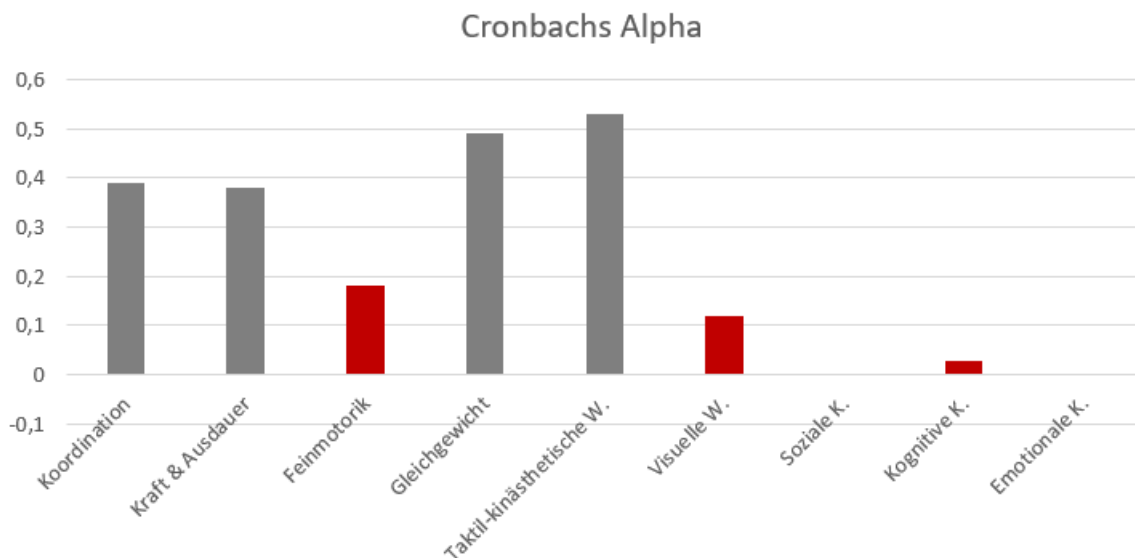


Abb. 4

Anmerkungen. Rot = sehr niedrig.

In Bezug auf die Itemschwierigkeit, lassen sich eindeutigeren Rückschlüsse ziehen. Ein Wertebereich von 0.2 bis 0.8 wird als angemessen erachtet, wobei der Bereich um den Wert 0.5 am besten ist (Döring & Bortz, 2016). Um den Wert 0.5 bewegen sich drei Items. Zum einen das Item „Wurf auf Zielscheibe“ (0.4), „rückwärts balancieren“ (0.48) und „Viereck formen“ (0.53). Das Item „Wasser umfüllen“ (0.84) befindet sich außerhalb des angemessenen Bereichs und wird demnach aus dem Erhebungsinstrument gestrichen. Alle anderen Items befinden sich zwar in dem angemessenen Bereich, jedoch zwischen den Werten 0.64 und 0.8, welche dafür stehen, dass die betroffenen Items sich im leichten Schwierigkeitsbereich befinden (Abb. 6).

Weitergehend kann auf die Itemtrennschärfe Bezug genommen werden. Hierbei stellen die Werte 0.3 bis 0.5 eine mittlere Trennschärfe und die Werte ab 0.5 eine hohe Trennschärfe dar (Döring & Bortz, 2016). Lediglich die Items „Gewichte sortieren“ (0.36), „Material ertasten“ (0.36), „Wandsitz“ (0.38) und „vorwärts balancieren“ (0.41) liegen im mittleren Bereich und das Item „rückwärts balancieren“ (0.55) im guten Bereich. Der Großteil der Items liegt im schlechten Bereich. Dennoch können Items nahe 0 aussortiert werden, da diese einzelnen Items nach den Werten mit dem restlichen Erhebungsinstrument kaum Überschneidungspunkte haben und demzufolge das Gesamtergebnis nicht angemessen repräsentieren. Nahe dem Wert 0 befinden sich die Items „Geduld“ (0), „Hilfsbereitschaft“ (0), „Viereck formen“ (0), „Labyrinth“ (0.02), „Bilderreihen“ (0.02), „Rasterbild erkennen“ (0.08), „Laufrichtung erkennen“ (0.08) und „Wasser umfüllen“ (0.1). Nach dieser Analyse müssten, wie bei der Cronbachs Alpha Analyse, die Kompetenzbereiche kognitive-, emotionale- sowie soziale Kompetenzen aussortiert werden. Zusätzlich müssten hierbei die visuellen Kompetenzen und die Feinmotorik aus dem Erhebungsinstrument gestrichen werden. Auch bei der Cronbachs Alpha Analyse schneidet der visuelle Kompetenzbereich und die Feinmotorik mit Werten von 0.12 und 0.18, nach den zuvor erwähnten Bereichen, am schlechtesten ab. Hierbei lassen sich Übereinstimmungen der Ergebnisse erkennen, welche die genannten Kompetenzbereiche als ungeeignet bewerten (Abb. 5).

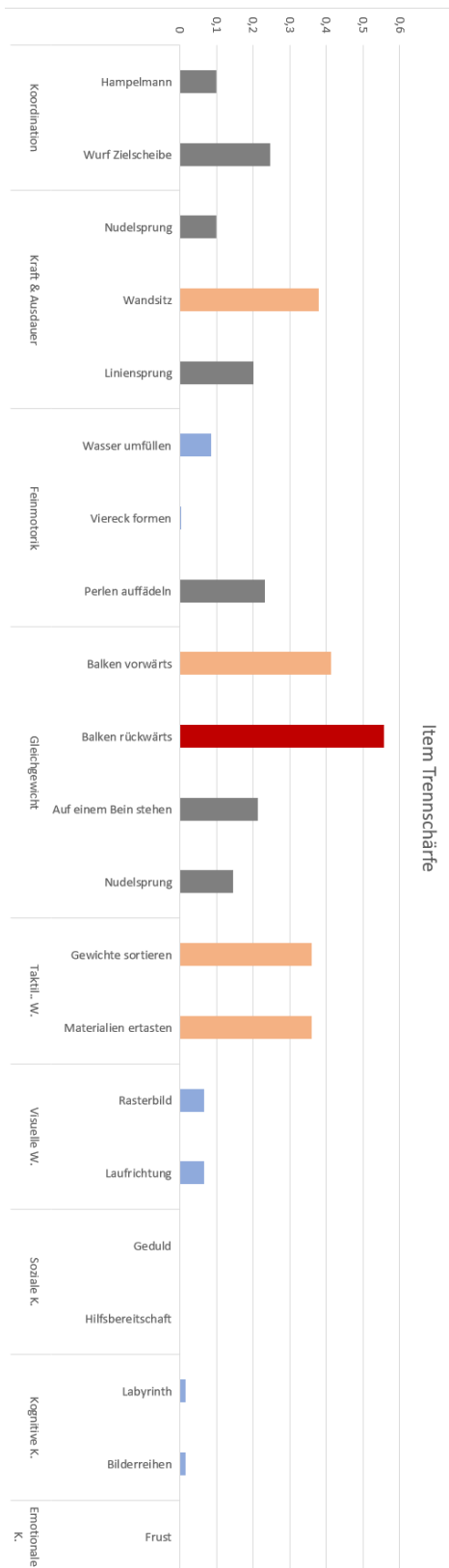


Abb. 5 Anmerkungen.
Rot= gut, Orange= mittel, Blau= nahe 0.

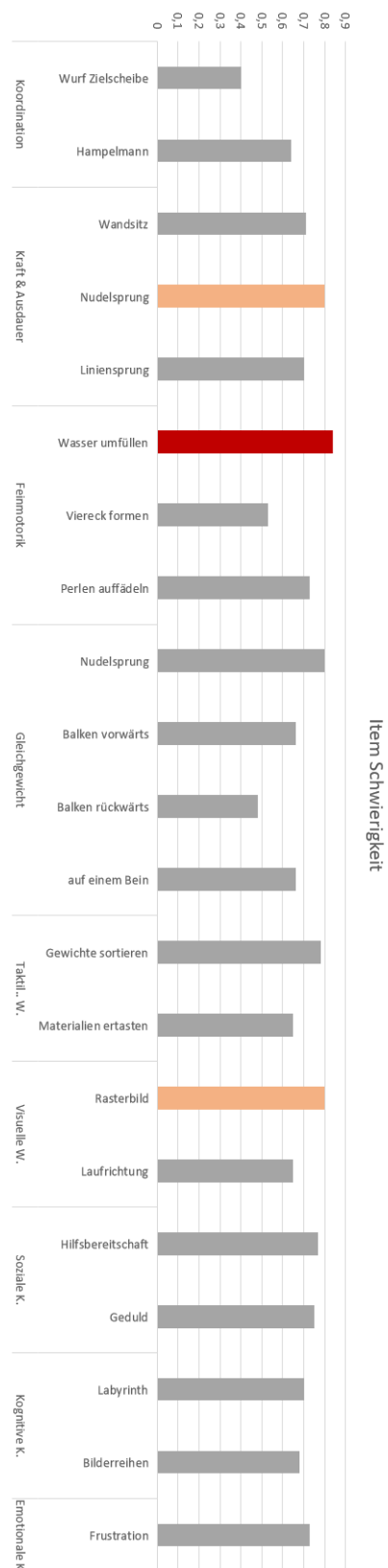


Abb. 6 Anmerkungen.
Rot= zu leicht, Orange= 0.8 Grenze.

Zugleich wurden die Mittelwerte der Kompetenzbereiche ermittelt. Bei gewissen Items schnitten die Schüler*innen durchschnittlich mit hohen Werten ab. Hierdurch wird ersichtlich, dass diese sich kaum bis gar nicht verbessern können und somit auch keine Entwicklung von Werten beobachtet werden konnte. Im Folgenden werden Items herausgestellt bei welchen die Schüler*innen durchschnittlich Werte gleich oder größer 3 erreicht haben.

Im Bereich der Feinmotorik betrifft dies das Item „Wasser umfüllen“ mit einem Wert von 3.37 (Abb. 7).

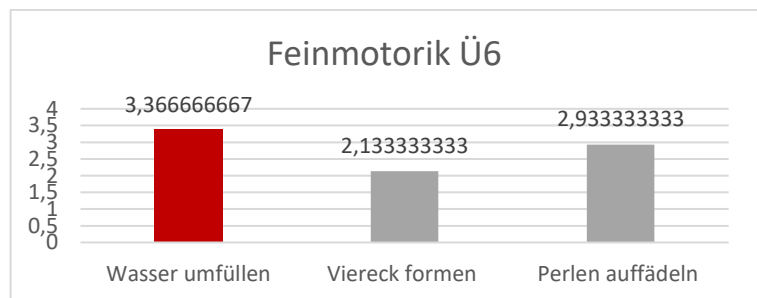


Abb. 7

Im Bereich Gleichgewicht und dementsprechend auch Kraft und Ausdauer, da hierbei eine Aufgabe für zwei Kompetenzbereiche genutzt wurde, betrifft es das Item „Nudelsprung“ mit einem Wert von 3.2 (Abb. 8 & 9).

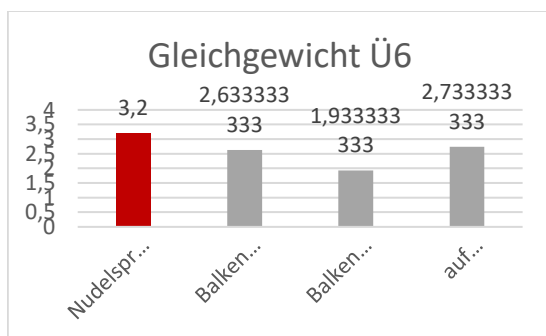


Abb. 8

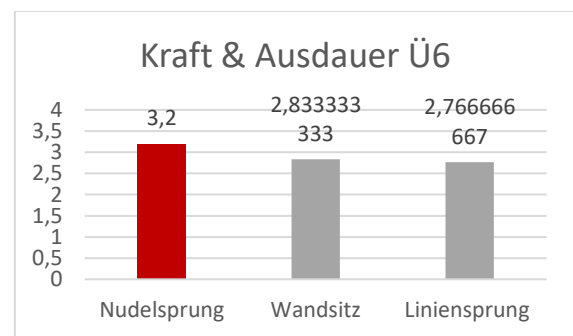


Abb. 9

Im Bereich der taktil-kinästhetischen Wahrnehmung überschreitet das Item „Gewichte sortieren“ die gesetzte Grenze von gleich oder höher 3 mit einem Wert von 3.13 (Abb. 10).

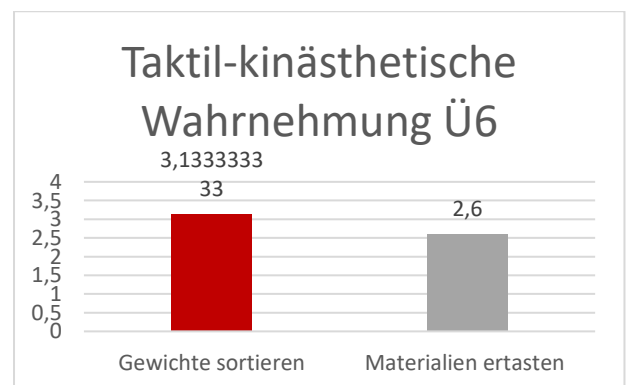


Abb. 10

Im Bereich der visuellen Kompetenzen betrifft es das Item „Rasterbild erkennen“ mit einem Wert von 3.17 (Abb. 11).

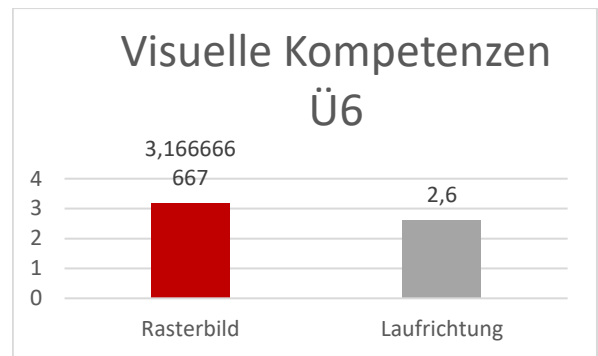


Abb. 11

Der Kompetenzbereich soziale Kompetenzen überschreitet mit beiden Items. Das Item „Hilfsbereitschaft“ hat einen Mittelwert von 3.07 und das Item „Geduld“ einen Wert von 3 (Abb. 12).

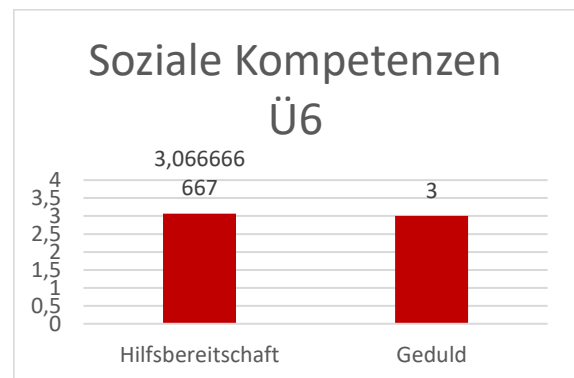


Abb. 12

Im Hinblick auf alle Ergebnisse müssten die Kompetenzbereiche „Soziale-, kognitive- und emotionale Kompetenzen“ aus dem Erhebungsinstrument entfernt werden. Da die sozialen-, kognitiven- sowie emotionale Kompetenzen den psychomotorischen Aspekt des Erhebungsinstruments ausmachen, sollten diese nicht entfernt, sondern überarbeitet beziehungsweise erneuert werden. Ebenso die „visuelle Kompetenzen“ und „Feinmotorik“ erwiesen sich als ungeeignet für das Erhebungsinstrument. Vor allem das Item „Wasser umfüllen“ aus dem Kompetenzbereich Feinmotorik, zeigte bei allen Analysekomponenten eine Schwäche. Demnach sollte es entfernt werden und in Zukunft möglicherweise durch ein neues Item ersetzt werden.

Innerhalb dieses Kapitels werden lediglich die Schwächen der Items aufgezeigt, um diese aussortieren zu können. Sinnig wäre es vereinzelt Items durch neue auszutauschen oder zu ergänzen. Folglich könnten sich auch die Cronbachs Alpha Werte verbessern, da diese bei einer größeren Anzahl von Items besser ausfallen. Die Ausarbeitung und Erprobung neuer Items überschreitet jedoch den Umfang dieses Projekts und müsste in einem Folgeprojekt erarbeitet werden. Bei dieser notwendigen Überarbeitung müssten ebenso neue Durchführungsmanuale erstellt werden, um die

verbesserte Adaption weiter prüfen zu können. Aufgrund der fehlenden Zeit innerhalb des Projekts konnten keinen neuen Items neu erstellt oder weiterentwickelt werden und demzufolge wurden auch keine neuen Durchführungsmanuale erstellt.

5.2 Ü9 Bereich

Nun wird auf die Überarbeitung des Ü9 Bereichs eingegangen. Auch hierbei werden Cronbachs Alpha, die Mittelwerte und Itemschwierigkeit sowie -trennschärfe beachtet. Im Gegensatz zu dem Ü6 Bereich liegen hierbei mehrere Erhebungszeitpunkte vor, zwischen welchen verglichen werden kann.

Auch hier kann kein Item nach der eigentlichen Differenzierung der Werte von Cronbachs Alpha aussortiert werden, da kein Kompetenzbereich einem brauchbaren Wert entspricht (Kuckartz et. al. 2013). Lediglich der Bereich taktil-kinästhetische Wahrnehmung grenzt mit einem Wert von 0,69 an der Grenze zum Wert 0,7. Auch im Ü9 Bereich enthalten die emotionalen Kompetenzen lediglich ein Item, weshalb Cronbachs Alpha nicht errechnet werden kann. Ebenso die Kognitiven Kompetenzen erlangen im Gegensatz zu den anderen Kompetenzbereichen einen deutlich niedrigeren Wert und könnten demnach aus dem Testverfahren gestrichen werden (Abb. 13).

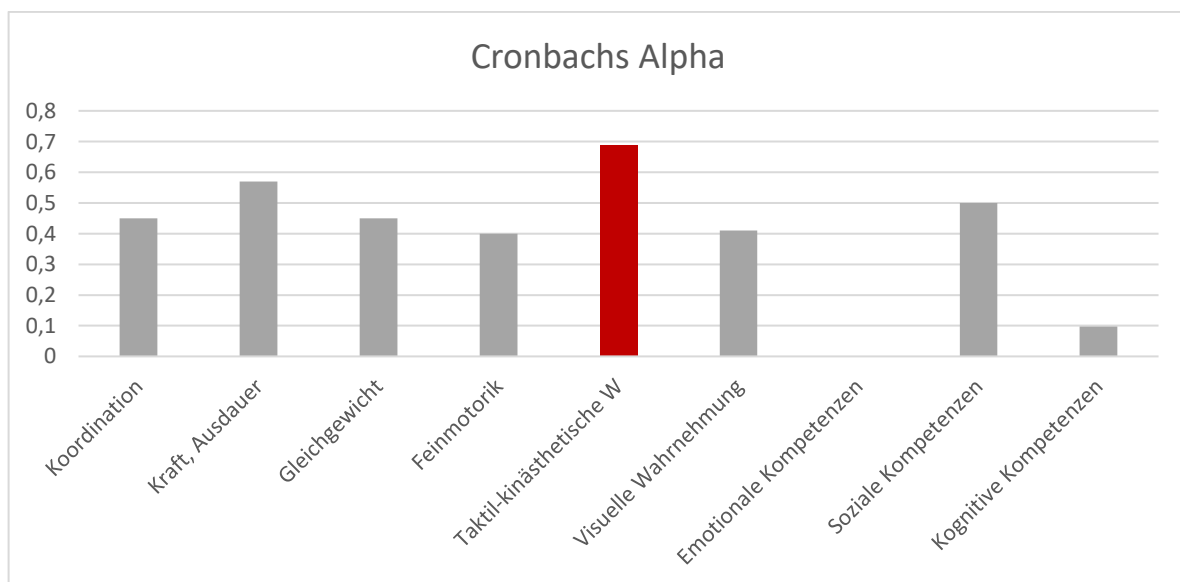


Abb. 13 Anmerkungen. Rot = Grenze 0.7 → 0.7-0.8 sind brauchbare Werte.

Bei der Itemtrennschärfe lässt sich herausfiltern, dass die Items „Wandsitz“, „Linien-sprung“, „Gewichte sortieren“ und „Material ertasten“ in einem guten Bereich sind. Ebenso befinden sich zehn Items im mittleren Bereich und repräsentieren demnach das Gesamtergebnis des Tests und messen das Zielkonstrukt akzeptabel. Die Items „Nudelsprung“, „Schleife binden“, „Labyrinth“ und „Bilderreihen“ haben Werte, die unter dem mittleren Bereich liegen und somit nicht angemessen das Gesamtergebnis repräsentieren. Demnach fällt der Kompetenzbereich kognitive Kompetenzen vollständig raus, aus dem Bereich Gleichgewicht zwei Items und aus dem Kompetenzbereich Kraft und Ausdauer ein Item (Abb. 14).

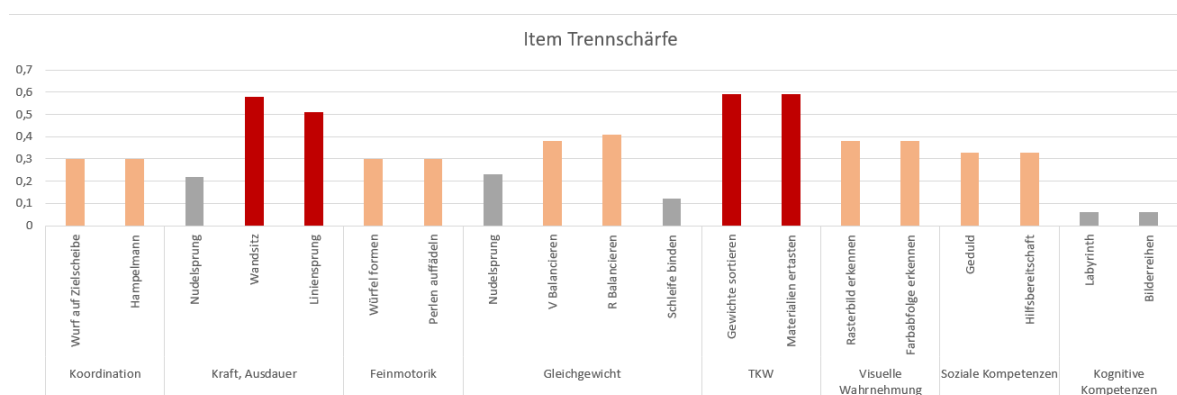


Abb. 14

Anmerkungen. Rot = gut, Orange = mittel.

Bei der Itemschwierigkeit wird deutlich, dass der Großteil der Items (13 von 20 Items) zu leicht ist. Die restlichen sieben Aufgaben kreisen mit Werten zwischen 0,46 und 0,74 um den optimalen Wert von 0,5. Die Differenzierung nach der Item Schwierigkeit, kann nicht vollständig zum Aussortieren der Items genutzt werden, da somit 65 Prozent der Items wegfallen würden und anschließend lediglich vier Kompetenzbereiche bestehen bleiben würden (Abb. 15).

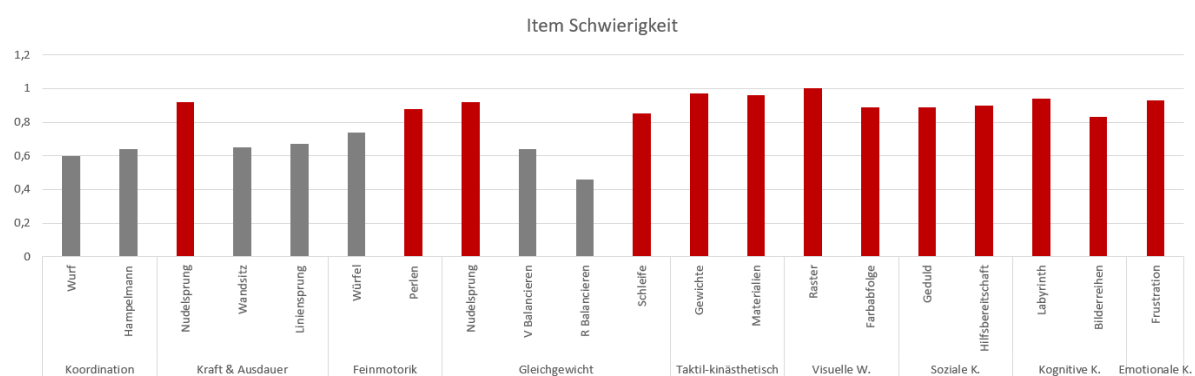


Abb. 15

Anmerkungen. Rot = zu leicht.

Folgend wird Bezug auf die Mittelwerte der einzelnen Aufgaben genommen. Zu jeder Aufgabe gibt es jeweils zwei bis drei Erhebungszeitpunkte, welche miteinander verglichen werden können. Bei den Kompetenzbereichen Emotionale Kompetenzen, Kognitive Kompetenzen und taktil-kinästhetische Wahrnehmung steigerten sich die Werte der teilnehmenden Schüler*innen, was darauf hinweist, dass sowohl ein Steigerungspotenzial als auch eine mögliche Auswirkung der Intervention möglich ist (Abb. 16, 17 & 18).

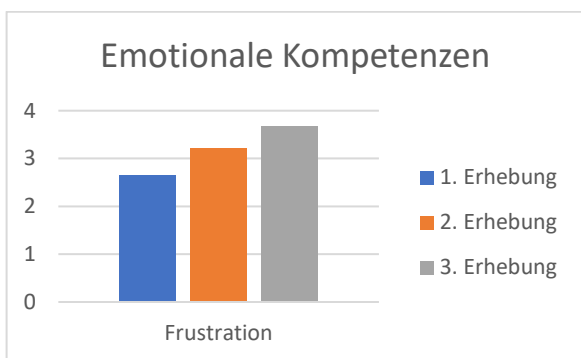


Abb. 16

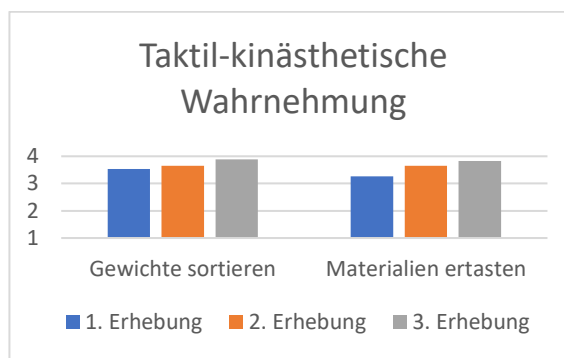


Abb. 17

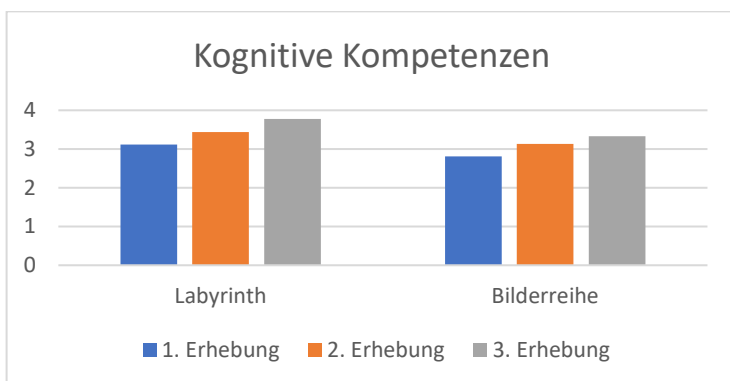


Abb. 18

Bei den Kompetenzbereichen Visuelle Wahrnehmung und soziale Kompetenzen, stiegen die Mittelwerte der Kinder bei einem Item kontinuierlich an und bei einem blieben sie stagnierend (Abb. 19 & 20).

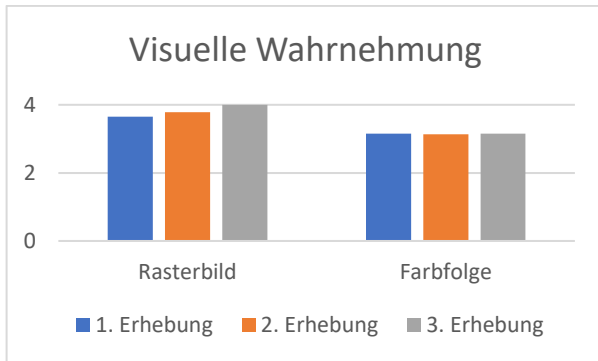


Abb. 19

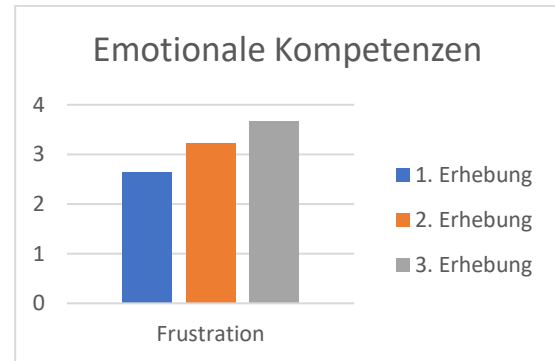


Abb. 20

Bei den Kompetenzbereichen Feinmotorik, Gleichgewicht und Kraft und Ausdauer wechselten die Items zwischen Anstieg, Abfall und Stagnation (Abb. 21, 22 & 23).

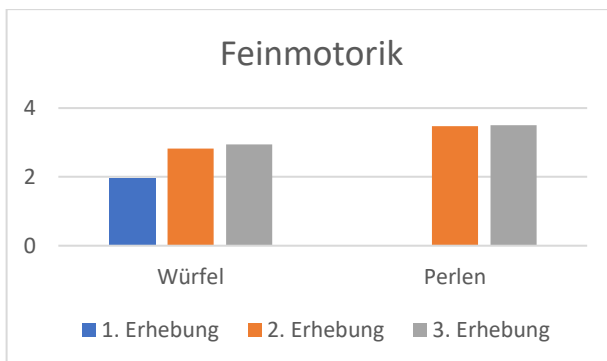


Abb. 21

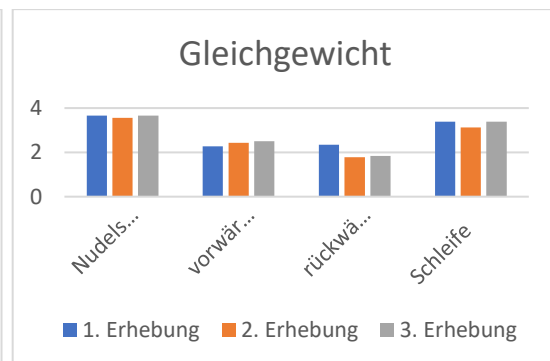


Abb. 22

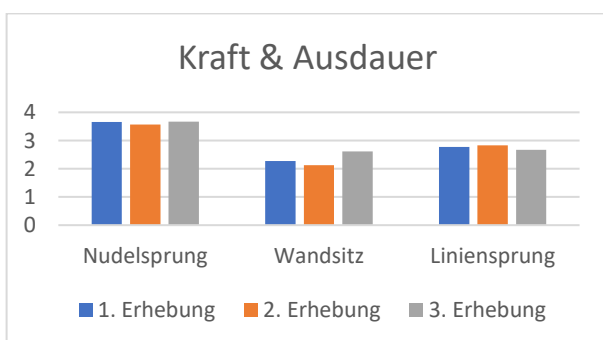


Abb. 23

In dem Kompetenzbereich Koordination war bei beiden Items ein Abfall der Werte zu erkennen (Abb. 24).

Wenn demnach die Mittelwerte als Indikator zum Aussortieren von Items genutzt werden, müssten all die Items, die einen Abfall oder eine Stagnation zu verzeichnen haben, aus dem Erhebungsinstrument gestrichen werden, da bei diesen kein Steigerungspotenzial und auch keine mögliche Auswirkung der Intervention zu erkennen ist.

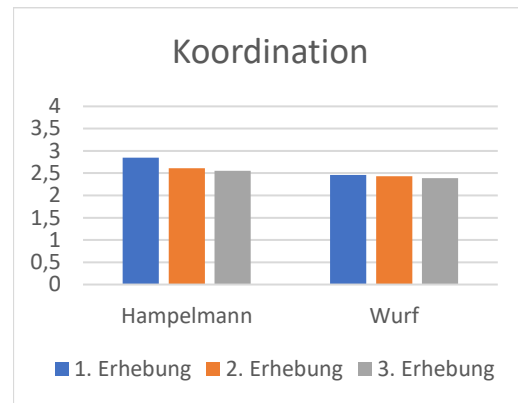


Abb. 24

Im Hinblick auf alle Ergebnisse muss der Kompetenzbereich emotionale Kompetenzen aufgrund eines einzelnen Items aus dem Erhebungsinstrument gestrichen werden. Ebenso die Kognitiven Kompetenzen, müssen nach der Analyse der Trennschärfe aufgrund der Leichtigkeit entfernt werden. Aus demselben Grund werden auch die Items „Nudelsprung“ (Gleichgewicht & Kraft & Ausdauer) und „Schleife binden“ (Gleichgewicht) bei der Überarbeitung des Manuals entfernt. Nach der Analyse der Mittelwerte wurde ersichtlich, dass sich die Schüler*innen im Kompetenzbereich Koordination von Erhebungszeitpunkt zu Erhebungszeitpunkt stetig verschlechtert haben. Auch im Ü9 Bereich wäre es sinnvoll vereinzelte Items durch neue auszutauschen oder zu ergänzen.

6. Diskussion

6.1 Beantwortung Forschungsfrage

Rückbindend an die Forschungsfrage „Inwiefern lässt sich der psychomotorische Test MotorikPlus von Zimmer auf andere Altersspannen adaptieren?“, lässt sich schlussfolgern, dass diese Frage nicht einfach beantwortet werden kann, da es sich hierbei um keine klare Antwort auf eine zuvor gestellte Hypothese handelt, sondern es um einen kontinuierlichen Prozess der Adaption eines Erhebungsinstruments geht. Dennoch lässt sich sagen, dass das vorliegende Forschungsprojekt wie geplant eine Grundlage für spätere Forschungsvorhaben bietet. Das Erhebungsinstrument muss folglich durch weitere Erhebungen verbessert werden. Erst nach erfolgreicher Normierung, welche zum jetzigen Zeitpunkt noch nicht erfolgt ist, kann von einem fertigen Erhebungsinstrument, beziehungsweise einer erfolgreichen Adaption eines bestehenden Erhebungsinstruments gesprochen werden. Grundsätzlich kann jedoch gesagt werden, dass es möglich ist ein bestehendes Erhebungsinstrument auf anderen Altersspannen zu adaptieren, wenn genügend Möglichkeiten zur Erprobung bestehen. Durch eine repräsentative Menge an kontinuierlichen Datenerhebungen lässt sich die Adaption weiter verbessern und letztendlich fertigstellen. Aufgrund der geringen Stichprobe innerhalb des Projekts und der aktuellen Covid-19 Pandemie wurde keine repräsentative Stichprobenmenge erzielt, woraufhin keine angemessene Adaption erfolgen konnte. Diese bedarf weiterer Überarbeitungsprozesse, welche den Projektrahmen zeitlich überschreiten.

6.2 Bewertung der Projektziele

Folgend wird Bezug auf die zuvor genannten Projektziele (Kapitel 3.2) genommen. Zu dem ersten Projektziel, bei welchem es darum ging, weitere Erhebungen zur Verbesserung des Manuals durchzuführen, kann gesagt werden, dass der Prozess nur teilweise umzusetzen war. Es fanden Erhebungen statt, welche zur Verbesserung beigetragen haben, dennoch nicht genug, um es finalisieren zu können. Ebenso konnte durch das Forschungsprojekt das Streben nach einer verbesserten Durchführungsqualität, sowie Praxis, realisiert werden. Durch kontinuierliche Nachbesprechungen unter den Studierenden wurde Abläufe verbessert und es kam eine gewisse Routine in die Handhabung der Aufgaben, wodurch die Durchführungsqualität

erheblich ansteigen konnte. Bezüglich der Messung von den psychomotorischen Kompetenzen konnten noch keine konkreten Ergebnisse erzielt werden, da zwar Vergleiche zwischen Schüler*innen und verschiedenen Erhebungszeitpunkten möglich waren, jedoch diese nur in einer geringen Menge erfolgten und es aufgrund einer hohen Fluktuation innerhalb der Klassen zu keinem stetigen Vergleich kommen konnte. Eine Auswertung der erhobenen Daten war mithilfe des Auswertungsprogramms RStudio möglich, jedoch brachte dies einige Schwierigkeiten mit sich, da aufgrund von fehlenden Informationen nachfolgende Fehler hätten verhindert werden können. Demzufolge konnten einige Items aufgrund fehlender Daten, nicht im Manual modifiziert sowie verbessert werden. Aufgrund der fehlenden Rahmeninformationen und Datenmenge, kamen unzureichende Werte bei der Auswertung heraus. Dies hatte zur Folge, dass sie aus dem Manual gestrichen, verbessert oder ausgetauscht werden mussten. Dieser Prozess war aufgrund der Menge an unzureichenden Ergebnissen zeitlich nicht mehr in das laufende Forschungsprojekt inkludierbar. Ein weiteres Projektziel bestand darin, die Lehrer*innen Perspektive innerhalb des Projekts, mithilfe eines zunächst leitfadengestützten Interviews zu erfassen. Dies wurde aufgrund des geringen zeitlichen Rahmens, innerhalb der Erhebungszeitpunkte, in eine schriftliche Befragung abgeändert. Hierdurch wurden tiefergehende Informationen erfasst, welche für Weiterentwicklung des Erhebungsinstruments erforderlich waren. Dennoch wäre ein Einblick in vielzählige Perspektiven, wie beispielsweise andere Interventionsklassen der teilnehmenden Schulen (Leseklasse Recklinghausen), wünschenswert gewesen. Diese Umsetzung scheiterte aufgrund der geringen Anzahl an kooperierenden Schulen. Ebenso sollte durch die schriftliche Befragung ein Einblick in die bewegten Pausen gewährt werden, um diese anschließend besser kontrollieren und verbessern zu können, damit ein möglichst großer Effekt durch diese Intervention erfolgen kann. Es wurde ein guter Einblick in die Interventionen ermöglicht, dennoch gelangen diese Informationen, aufgrund einer verspäteten Kommunikation seitens der Kooperationsschulen, erst zum Ende des Projekts zu der Projektgruppe, wodurch eine Unterstützung bei der Auswahl der Übungen und eine stetige Anleitung nicht gewährleistet werden konnte.

6.3 Kritische Reflexion der Projektergebnisse

Es folgt eine kurze Zusammenfassung der Projektergebnisse. Es wurde das psychomotorische Beobachtungsverfahren „MotorikPlus“ von Zimmer (2021) auf andere Altersspannen adaptiert. Dafür wurden Items aus verschiedenen motorischen Testverfahren hinzugefügt und einige herausgestrichen. Zusätzlich wurden entweder die Bewertungsskalen von bereits vorhandenen Items des „MotorikPlus“ verändert oder ebendiese Items abgeändert. Es wurde ein Testmanual und ein Auswertungsbogen erstellt und während drei Erhebungszeitpunkten eingesetzt und überarbeitet. Nach der quantitativen Untersuchung der Items wurde festgestellt, dass die meisten Items unbrauchbar sind (siehe Kap. 5). Für die Altersspanne Ü6 weisen wenige Items eine angemessene Itemschwierigkeit auf und der Großteil der Items hat eine schlechte Itemtrennschärfe (siehe Kap. 5). Für die Altersspanne Ü9 sind die Itemtrennschärfen besser als bei Ü6 (siehe Kap. 5). Aber auch hier sind wenige Items angemessen schwer (siehe Kap. 5). Es müssten mehrere Kompetenzbereiche überarbeitet sowie ein paar einzelne Items aus den Kompetenzbereichen herausgenommen werden. Mit wenigen Ausnahmen sind die Reliabilitätsmessungen insgesamt als mangelhaft einzustufen. Die Durchführungsobjektivität ist teilweise durch ein Anleitungsmanskript im Testmanual gegeben. Jedoch kam es dennoch hin und wieder zu Verständnisproblemen der Schüler*innen während der Testanweisung und dadurch wurde die Durchführungsobjektivität eingeschränkt. Auch wurden die standardisierten Testanleitungen sprachlich unterschiedlich betont, was die Durchführungsobjektivität ebenfalls einschränkte (siehe Kap. 4.1). Die Auswertungsobjektivität ist durch eine umfassende Beschreibung der möglichen Antworten der Schüler*innen und deren Bewertung im Testmanual gegeben. Das Testmanual des MotorikPlus von Zimmer (2021) wurde verwendet und adaptiert, welches hinsichtlich der Inhalte, Struktur und Praxis-tauglichkeit als valide gilt und daher der inhaltlichen Validität entspricht.

Der „MotorikPlus“ konnte auf eine andere Altersspanne adaptiert werden, auch wenn das Verfahren aufgrund von mangelhafter Reliabilität noch überarbeitet werden muss.

6.4 Limitation der Forschung

Das Projekt wurde im Verlauf durch verschiedene Aspekte limitiert. Ein ausschlaggebender Aspekt ist die Corona-Pandemie, welche den Projektverlauf stark beeinflusste, da Datenerhebungen in den Schulen folglich nur eingeschränkt möglich waren. So konnte in Recklinghausen, innerhalb der Altersspanne Ü6, nur eine Erhebung stattfinden. Bei Betrachtung der Rückmeldung von der Bewegungsklasse in Recklinghausen (Anhang A) fällt auf, dass diese Einschränkung enorme Auswirkungen auf die Auswertungsmöglichkeiten des Ü6er Bereichs hat. Die Bewegungsklasse in Recklinghausen verfolgte die Intervention „Bewegung“ sehr gewissenhaft, weswegen eine zweite Erhebung von signifikantem Vorteil gewesen wäre. Insbesondere im Hinblick darauf, dass die Bewegungsklasse der Altersspanne Ü9 nur eingeschränkt die Anforderungen an die Bewegungspausen erfüllte. Dass kein Termin für eine weitere Erhebung in Recklinghausen stattfinden konnte, lag an der Pandemie und dem dadurch bedingten Wunsch nach Kontaktvermeidung seitens der Schule, jedoch auch an mangelnder Kommunikation. Limitiert wird die Aussagekraft der Ergebnisse ebenfalls durch die kleine Stichprobe. Die Stichprobe ist als nicht repräsentativ zu beschreiben, wodurch dementsprechend keine Normierung stattfinden kann. Daher sollte das Manual in einer repräsentativen Stichprobe weiter erprobt werden, insbesondere die Skalenentwürfe müssen überprüft werden. Ein wichtiger limitierender Faktor ist zudem die Gruppengröße der Projektgruppe, welche innerhalb des Projektjahres auf fünf Studierende geschrumpft ist. Dadurch wurden die Möglichkeiten innerhalb des Projekts stark eingeschränkt, sodass ein Fokus gesetzt werden musste. Dadurch konzentriert sich die Forschung insbesondere auf die quantitative Auswertung der Reliabilitätskoeffizienten.

6.5 Forschungsbedarf

Das vorliegende Erhebungsinstrument bietet eine Grundlage für spätere Forschungsvorhaben und ist demgemäß eine Pilotstudie. Forschungs- sowie Verbesserungsbedarf sind auch nach Abschluss der Messergebnisse weiterhin vorhanden, da das Erhebungsinstrument durch weitere Erhebungen verbessert werden muss, um es normieren zu können. Sowie bereits benannt kann erst nach erfolgreicher Normierung

von einem fertigen Erhebungsinstrument, beziehungsweise einer erfolgreichen Adaption eines bestehenden Erhebungsinstruments gesprochen werden.

Der Zusammenhang mit der Adaptierung der Items ergibt außerdem, dass es sich bei dem MotorikPlus um einen alltagstauglichen Test handelt, da viele Items dafür geeignet sind, diese im Kindergartenalltag zu beobachten (Zimmer, 2021). Während dies wiederum nicht mit dem selbstentwickelten Testinstrument gleichzusetzen ist, da die vorhandenen Aufgaben in dem Erhebungsinstrument kaum in den Schulalltag integriert werden können. In Anbetracht der bisher statischen Lehrform innerhalb von Schulen, ist zu beachten, dass wenig Handlungsfreiheiten zur Verfügung stehen, die Items innerhalb des Schulalltags durchzuführen, um die Kompetenzen der Kinder zu beobachten und zu dokumentieren. Überdies stellt das Forschungsprojekt eine explorative Voruntersuchung dar, die zur Theorie- und Hypothesenbildung anregt. Aufgrund der ungleichen Verteilung der teilnehmenden Schüler*innen lassen sich keine geschlechterspezifischen Unterschiede feststellen. Daher besteht durch weitere Entwicklung sowie Verbesserung des Erhebungsinstruments, und vergrößerter Stichprobe, eine gewinnbringende Chance das Erhebungsinstrument für weitere Forschungsfragen, wie beispielsweise unsere erste entworfene Fragestellung zu nutzen. Dabei muss beachtet werden, dass die äußerlichen gesellschaftlichen Veränderungen, wie beispielsweise die Covid-19-Pandemie, immer gegeben sein können und somit das Forschungsvorhaben beeinflussen.

6.6 Kritische Reflexion Projektverlauf & Projektorganisation

Im Hinblick auf die Bewertung der Projektziele kann festgehalten werden, dass die Projektgruppe innerhalb der zweiten Hälfte des Projektzeitraums die zuvor gesetzten Meilensteine erreichen konnte. Es konnten weitere Daten, mittels des überarbeiteten Manuals, erhoben und diese anschließend ausgewertet sowie interpretiert werden. Insbesondere die Auswertung stellt ein großes Ziel der zweiten Hälfte dar. Auch die E-Learning Aufgaben konnten erledigt und eingereicht werden.

Bei rückblickender Betrachtung des Projektverlaufs fällt auf, dass die zweite Hälfte des Projektjahres sehr kurz war, da es nicht das gesamte Sommersemester, sondern nur drei Monate des Semesters beinhaltet. Außerdem befinden sich in diesem Zeitraum viele Feiertage, weswegen die Zeit ebenfalls gekürzt wurde. Insbesondere im

Hinblick auf die Zusammenarbeit mit den Schulen fiel die kurze Zeitspanne auf. Hinzukommend war, dass die Ferienzeiten der Schulen berücksichtigt werden mussten, wodurch sich terminliche Möglichkeiten stark eliminierten. Weiter beeinflusste die Corona-Pandemie die Terminfindung mit den Schulen. Hierbei ist jedoch als positiv zu verzeichnen, dass die Projektgruppe bereits in der ersten Hälfte einige Datenerhebungen durchführen konnte und somit bereits eine Auswertungsgrundlage vorzuweisen hat. Dadurch wurde bereits in der ersten Hälfte die Durchführung des Forschungsvorhabens sichergestellt.

Bezüglich der Auswertung kann angeführt werden, dass rückblickend eine frühzeitigere Auseinandersetzung mit Auswertungsmethoden und -bedingungen sinnvoll gewesen wäre. So war den Studierenden anfangs nicht bewusst, dass mehr als ein Item je Skala vorhanden sein muss, damit beispielsweise Cronbachs Alpha berechnet werden kann. Die Studierenden wurden darauf hingewiesen, haben dies allerdings in einem anderen Kontext eingeordnet. Auch wäre eine frühere Auswertung, beispielsweise nach den Erhebungszeitpunkten, sinnvoll gewesen, um direkt adäquat darauf reagieren zu können und die Adaption für weitere Erhebungszeitpunkte anzupassen.

Zu Beginn des Projekts verteilten die Gruppenmitglieder spezifische Rollen, anhand der individuellen Stärken und Schwächen. Diese Aufteilung der Verantwortlichkeiten funktionierte in der ersten Hälfte sehr gut und wurde auch in der zweiten Hälfte beibehalten. Die Projektgruppe verteilte auch in diesem Projektabschnitt Schwerpunkte und Aufgaben, sodass zielorientiert und strukturiert gearbeitet werden konnte. Durch die individuellen Verantwortlichkeiten gelang es sehr gut Termine, Absprachen und Aufgaben einzuhalten und zu bearbeiten. Hier ist die minimierte Gruppengröße zu erwähnen, da die Projektgruppe fünf Studierende zählt. Dadurch mussten innerhalb des Projekts Schwerpunkte gesetzt werden, wodurch sich insbesondere mit der quantitativen Auswertung und den Gütekriterien beschäftigt wurde. Die Zusammenarbeit innerhalb der kleinen Gruppe gelang sehr gut.

6.7 Rückblick

Mithilfe der qualitativen schriftlichen Befragung wurden zusätzliche Informationen gesammelt. Aus diesen konnten Rückschlüsse gezogen und somit weitergehend reflektiert werden. Zum einen wurden positives Feedback, sowie Verbesserungsvorschläge

von den Lehrer*innen gegeben. Zusätzlich verdeutlichte dies, wie die Eltern das Verhalten der Kinder nach den Interventionen aufgenommen haben. Zu den positiven Aspekten zählt zum einen, dass in der Leseklasse die Lesezeiten von den Kindern positiv angenommen und wahrgenommen wurden (Anhang A, Z. 15 f) sowie die Lesesintervention zum Runterkommen genutzt werden konnten (Anhang A, Z. 23). Zum anderen teilten die Eltern den Lehrer*innen mit, dass ihre Kinder häufig gut gelaunt nach Hause kamen und von den Geschichten erzählten (Anhang A, Z. 27f). In der Bewegungsklasse in Gelsenkirchen fiel auf, dass die Bewegung zur positiven Verhaltensänderung geführt hat (Anhang A, Z. 33f). Recklinghausen gab ebenfalls an, dass die Kinder nach den Bewegungsphasen viel motivierter und konzentrierter bei der Arbeit waren (Anhang A, Z. 61) und sich der Zusammenhalt eher verstärkt hat (Anhang A, Z.64). Ebenso hatten alle Kinder Spaß und zeigten eine große Begeisterung bezüglich der Interventionen (Anhang A Z. 68, 74).

Auch gaben die Lehrer*innen der Bewegungsklassen an, dass sie mehr Unterstützung und Anleitung benötigt sowie feste Termine zur Überprüfung gebraucht hätten (Anhang A, Z. 47f, 71). In den Schulen wurden Rahmenbedingungen vorgegeben, diese beinhalten, dass die Kinder sich mindestens zwei mal fünf Minuten an ihrem Stuhl innerhalb der Klasse bewegen sollen. Zu Beginn des Projektes wurden diesbezüglich grobe Konzeptideen besprochen, beispielsweise, dass eine Klasse Bewegung anhand von YouTube Videos durchführt, und ebenso wurde den Klassen eine mündliche Unterstützung zugesprochen, welche sie jedoch nicht im vollen Maße nutzen konnten. Für die Zukunft könnte innerhalb des Projekts darauf geachtet werden, dass klare Konzepte besprochen und vorgelegt werden müssen sowie diese im Verlauf des Projektes durch mündliches Feedback überprüft werden. Durch einzelne Aussagen der Kinder bestätigte sich die Annahmen, dass die Lehrer*innen mit den Bewegungsinterventionen überfordert waren und diese somit nicht kontinuierlich durchgeführt wurden. Dies hätte mit regelmäßigen Überprüfungsterminen unterbunden werden können, um ebenso die Studienergebnisse nicht zu verfälschen. Mit veränderten Rahmenbedingungen würden die Lehrer*innen der Bewegungsklasse erneut an der Studie teilnehmen und die Bewegungen in ihrem Klassealltag integrieren (Anhang A, Z. 47f, 71). Im Gegensatz zu den Bewegungsklassen war die Umsetzung der Interventionen in der Leseklasse alltagstauglicher, weswegen sie wiederholt an

der Studie teilnehmen würden (Anhang A, Z. 19f). Nach eigenen Aussagen nahmen sie positive Aspekte bei ihren Schüler*innen wahr und die Lesepausen ließen sich gut in der Klassealltag integrieren.

Ein Aspekt, der sich negativ auf das Projektgeschehen auswirkte, war, dass einige Verhaltensveränderungen seitens der Lehrer*innen beobachtet wurden. In der schriftlichen Befragung gaben die Bewegungsklassen an, dass es unmittelbar nach der Bewegungseinheit zu einer motorischen Unruhe kam (Anhang A Z.32f). Dieser äußerte sich in einem gegenseitigen aufhetzen oder einem albernen Verhalten, welches sich ebenso in der Datenerhebung zeigte. Einige der Kinder ließen sich schnell ablenken oder verglichen sich untereinander, wodurch ein Konkurrenzkampf entstand. Folglich gab es eine angespannte Atmosphäre und es kam dazu, dass die Kinder ihre Werte der Aufgaben miteinander verglichen und diese bewerteten. In Bezug auf die Vergleiche unter den Kindern kam es vereinzelt zu Aussagen wie „Fließt das hier in meine Note mit ein?“. Demzufolge wurde ersichtlich, dass es sinnvoll wäre zu Beginn des Projektes deutlich zu machen, dass keine der Aufgaben in die Note der Kinder miteinfließt und sie alles im freiwilligen Rahmen machen und zu nichts davon gezwungen werden. Zusätzlich zeigten einige Kinder ein Gefühl von Scham, welches die Ausführung der Aufgaben erschwerte. Es fielen Aussagen wie „Ich möchte das nicht machen, die anderen Kinder gucken zu“ und „Ich kann das nicht“. Hierdurch wurde deutlich, dass nicht alle Aufgaben für die Kinder gleich bewältigbar und ausführbar waren.

Im gesamten meldeten die Klassen mit dem Förderschwerpunkt emotionale und soziale Entwicklung zurück, dass die Einstellung der Kinder, gegenüber dem Projekt und den Interventionen, sehr zwiespalten waren. Zum einen wurde es positiv von ihnen aufgenommen und zum anderen mussten Kinder unterstützend von den Lehrer*innen und oder Integrationshelfer*innen motiviert und begleitet werden (Anhang A, Z.52 f). Dass im Gegensatz zu den anderen Schulen, vermehrt negatives Feedback aus Gelsenkirchen stammt, könnte mit dem durchschnittlichen Alter der Kinder zusammenhängen und auch eventuell auf den Förderbedarf zurückzuführen sein. Die Schüler*innen im Alter von 13 bis 15 Jahren, waren sich „zu cool“ (Anhang A, Z.23 ff) für die Aufgaben, während die Kinder aus Recklinghausen, im Alter von 7 bis 9 Jahren, keiner solcher Aussagen tätigten und mit Begeisterung teilnahmen (Anhang A, Z.68, 74).

7. Fazit

Als Fazit kann gezogen werden, dass es durchaus möglich ist ein bestehendes Erhebungsinstrument auf eine andere Altersspanne zu adaptieren. Erfolgen kann dies durch das Hinzufügen und Herausnehmen von Items sowie die Modifizierung von bereits bestehenden Items. Aufgrund verschiedener limitierender Faktoren, wie die Covid-19-Pandemie und die Gruppengröße konnte die Adaption nicht im Rahmen des Projekts fertiggestellt werden. Die Reliabilitätsüberprüfung zeigte, dass der Großteil der Items unzureichender Güte ist, sodass eine Überarbeitung notwendig ist. Damit das Erhebungsinstrument in einem nächsten Schritt normiert werden kann, müssen die Items der momentanen Adaption überarbeitet werden und an einer repräsentativen Stichprobe erprobt werden. Noch stellt die Stichprobe keine Repräsentation der Teilpopulation, von Schüler*innen im Alter von 6-15, dar. Dennoch konnte eine Grundlage für weitere Forschungsvorhaben gesetzt werden. Die Adaption des Beobachtungsinstruments „MotorikPlus“ von Zimmer (2021) bleibt realistisch.

8. Literaturverzeichnis

- Blanz, M. (2021). *Forschungsmethoden und Statistik für die Soziale Arbeit – Grundlagen und Anwendungen* (2. Aufl.). Stuttgart: Kohlhammer.
- Bortz, J., & Döring, N. (2006). *Qualitative Methoden*. In *Forschungsmethoden und Evaluation* (pp. 295-350). Berlin/ Heidelberg: Springer.
- Bundesministerium für Familie, Senioren, Frauen und Jugend [BMFSFJ] (2022). *Aktionsprogramm „Aufholen nach Corona für Kinder und Jugendliche“*. Verfügbar unter: <https://www.bmfsfj.de/bmfsfj/themen/corona-pandemie/aufholen-nach-corona>, zuletzt geprüft am 13.06.2022.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (5. überarbeitete, aktualisierte und erweiterte Auflage). Berlin/Heidelberg: Springer.
- Döring, N., & Bortz, J. (2016). *Stichprobenziehung*. In *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (S. 291-319). Berlin/ Heidelberg: Springer.
- Fisseni, H. (1990). *Lehrbuch der psychologischen Diagnostik*. Göttingen: Hofgrefe.
- Gerhold, L., Holtmannspötter, D., Neuhaus, C., Schüll, E., Schulz-Montag, B., Steinmüller, K., & Zweck, A. (Hrsg.). (2015). *Standards und Gütekriterien der Zukunftsforschung: Ein Handbuch für Wissenschaft und Praxis*. Springer VS.
- Hug, T., Poscheschnik, G. (2020). *Empirisch forschen: Die Planung und Umsetzung von Projekten im Studium* (3. überarb. und ergänzte Aufl.). UVK Verlag.
- Kelava, A., & Moosbrugger, H. (2020). *Deskriptivstatistische Itemanalyse und Testwertbestimmung*. In H. Moosbrugger & A. Kelava (Hrsg.) *Testtheorie und Fragebogenkonstruktion* (3. Aufl.) (pp. 143-158). Heidelberg: Springer.
- Krumm, S., Schmidt-Atzert, L. & Amelang, M. (2021). *Grundlagen diagnostischer Verfahren*. In L. Schmidt-Atzert, S. Krumm, M. Amelang (Hrsg.). *Psychologische Diagnostik* (6. Aufl.) (S.39-208). Heidelberg: Springer.
- Kuckartz, U., Rädiker, S., Ebert, T. & Schehl, J. (Hrsg.). (2013). *Statistik - Eine verständliche Einführung* (2. Aufl.). Wiesbaden: Springer Fachmedien.

- Loy, T. (2015). *SPSS oder R?*. Statistik und Beratung – Daniela Keller. Verfügbar unter: <https://statistik-und-beratung.de/2015/02/spss-oder-r/>, zuletzt geprüft am 16.06.2022.
- Lüdders, L. (2016). *Fragebogen- und Leitfadenskonstruktion*. Ein Handbuch für Studium und Berufspraxis. Bremen: Apollon.
- Moosbrugger, H., & Kelava, A. (2020). *Qualitätsanforderungen an Tests und Fragebogen („Gütekriterien“)*. In H. Moosbrugger & A. Kelava (Hrsg.) *Testtheorie und Fragebogenkonstruktion* (3. Aufl.) (S. 13-38). Heidelberg: Springer.
- Moosbrugger, H., Gäde, J. Schermelleh-Engel, K. & Rauch, W. (2020). *Klassische Testtheorie*. In H. Moosbrugger & A. Kelava (Hrsg.) *Testtheorie und Fragebogenkonstruktion* (3. Aufl.) (S.275-304). Heidelberg: Springer.
- Moosbrugger, H., Schermelleh-Engel, K., Gäde, J. & Kelava, A. (2020). *Testtheorien im Überblick*. In H. Moosbrugger & A. Kelava (Hrsg.) *Testtheorie und Fragebogenkonstruktion* (3. Aufl.) (S.251-273). Heidelberg: Springer.
- Pfeiffer, D. K., & Püttmann, C. (2018). *Methoden empirischer Forschung in der Erziehungswissenschaft: Eine Einführung* (6. überarb. und erweiterte Aufl.). Münster/NewYork: Waxmann.
- Roch, S. (2017). *Der Mixed-Methods-Ansatz*. Forschendes Lernen an der Europa-Universität Flensburg–30 Erhebungsmethoden, 95-110.
- Spektrum. (n.d.). *Lexikon der Psychologie. Pilotstudie*. Heidelberg: Spektrum. Verfügbar unter: <https://www.spektrum.de/lexikon/psychologie/pilotstudie/11578>, zuletzt geprüft am 10.06.2022.
- Zimmer, R. (2021). *MotorikPlus: Beobachtung motorischer, sensorischer, emotionaler, sozialer und kognitiver Kompetenzen von Kindern im Alltag von Kindertageseinrichtungen: Manual*. Freiburg im Breisgau: Herder Verlag.
- Züll, C., & Menold, N. (2014). *Offene Fragen*. In *Handbuch Methoden der empirischen Sozialforschung* (S. 713-719). Wiesbaden: Springer VS.

9. Abbildungsverzeichnis

1. Tabelle: Itemschwierigkeiten.
2. Tabelle: Cronbachs Alpha.
3. Tabelle: korrigierte Itemtrennschärfe.
4. Cronbachs Alpha. Ü6 Bereich.
5. Itemtrennschärfe. Ü6 Bereich.
6. Itemschwierigkeit. Ü6 Bereich.
7. Mittelwerte. Kompetenzbereich Feinmotorik. Ü6 Bereich.
8. Mittelwerte. Kompetenzbereich Gleichgewicht. Ü6 Bereich.
9. Mittelwerte. Kompetenzbereich Kraft & Ausdauer. Ü6 Bereich.
10. Mittelwerte. Kompetenzbereich Taktil-kinästhetische Wahrnehmung. Ü6 Bereich.
11. Mittelwerte. Kompetenzbereich Visuelle Kompetenzen. Ü6 Bereich.
12. Mittelwerte. Kompetenzbereich Soziale Kompetenzen. Ü6 Bereich.
13. Cronbachs Alpha. Ü9 Bereich.
14. Itemtrennschärfe. Ü9 Bereich.
15. Itemschwierigkeit. Ü9 Bereich.
16. Mittelwerte. Kompetenzbereich Emotionale Kompetenzen. Ü9 Bereich.
17. Mittelwerte. Kompetenzbereich Taktil-kinästhetische Wahrnehmung. Ü9 Bereich.
18. Mittelwerte. Kompetenzbereich Kognitive Kompetenzen. Ü9 Bereich.
19. Mittelwerte. Kompetenzbereich Visuelle Wahrnehmung. Ü9 Bereich.
20. Mittelwerte. Kompetenzbereich Emotionale Kompetenzen. Ü9 Bereich.
21. Mittelwerte. Kompetenzbereich Feinmotorik. Ü9 Bereich.
22. Mittelwerte. Kompetenzbereich Gleichgewicht. Ü9 Bereich.
23. Mittelwerte. Kompetenzbereich Kraft & Ausdauer. Ü9 Bereich.
24. Mittelwerte. Kompetenzbereich Koordination. Ü9 Bereich.

10. Anhang

Anhang A

1 **Transkript I**

2 Leseklasse Gelsenkirchen

3 *Welchen Effekt haben Sie durch die jeweiligen Interventionen gespürt?*

4 Einige Schüler haben die Zeit für sich genutzt, um im Schulalltag etwas
5 „runterzukommen“ und die letzten zwei Stunden mit neuer Ruhe anzugehen (gelesen
6 wurde nach der 2. Pause)

7 *Hat sich etwas im Klassengefüge verändert? Wenn Ja, was hat sich verändert?*

8 Es hat sich für uns als Lehrer nochmals gezeigt, welche Schüler schon „etwas zu
9 cool“ für solche Angebote sind und welche dies gut nutzen können. Am
10 Klassengefüge hat sich nichts verändert.

11 *Welche Intervention wurde durchgeführt und in welchem Maße?*

12 Lesen diverser Bücher, wenn möglich täglich ein Kapitel für circa 15 bis 20 Min.

13 „Das fliegende Klassenzimmer“, „Drei Fragezeichen Romane“ & „100-Gefahren-
14 Geschichten“.

15 *Gab es ein Feedback von den Eltern? Wenn Ja, welches Feedback?*

16 Es gab kein spezielles Feedback. Zwei Eltern nahmen die Lesezeit als positiv für ihre
17 Kinder wahr (bedingungslose Belohnung).

18 *Würden Sie an der Studie noch mal teilnehmen wollen? Welche Gründe spielen bei
19 der Entscheidung eine Rolle?*

20 Als Lesegruppe würde ich wieder teilnehmen, weil es sich gut in den Alltag
21 integrieren lässt und für eine Schüler positive Effekte hat.

22 *Wie wurden die Interventionen von den Kindern angenommen? (Hat den Kindern es
23 Spaß gemacht oder waren sie abgeneigt?)*

24 Einige Schüler konnten es zum „Runterkommen“ gut annehmen, andere waren
25 schon etwas „zu cool“ und haben sich teilweise rausgenommen und Auszeiten auf
26 dem Flur gemacht.

27 *Welche Veränderung haben die Eltern an Ihren Kindern bemerkt?*

28 Von zwei Eltern kam die Rückmeldung, dass ihre Kinder häufig gut gelaunt nach
29 Hausen kamen und von der Geschichte erzählt, haben.

30

31 Bewegungsklasse Gelsenkirchen

32 *Welchen Effekt haben Sie durch die jeweiligen Interventionen gespürt?*

- 33 Je nach Tagesform/ Klassenkonstellation hat die Bewegung, zu einer erhöhten
34 motorischen Unruhe (vgl. Albernheit, gegenseitiges Aufpushen) oder zur positiven
35 Verhaltensänderung (vgl. beruhigend; verringerte motorische Unruhe) geführt.
- 36 *Hat sich etwas im Klassengefüge verändert? Wenn Ja, was hat sich verändert?*
- 37 Hohe Fluktuation der Klasse (vgl. Zu- & Wegzüge). Von der ursprünglichen
38 Lerngruppe sind lediglich 5 Schüler noch in der Klasse.
- 39 *Welche Intervention wurde durchgeführt und in welchem Maße?*
- 40 Angeleitete Rennspiele vor Schulbeginn (circa 15 min) und ein abwechselndes
41 Angebot. Verlängerte Bewegungspausen mit Ballspielangeboten (2-mal circa 10
42 min). Kurze Bewegungsangebote (Yoga, Stretching → angeleitet) während des
43 Unterrichts, bei Phasenwechsel → circa 15 min insgesamt.
- 44 *Gab es ein Feedback von den Eltern? Wenn Ja, welches Feedback?*
- 45 Nein.
- 46 *Würden Sie an der Studie noch mal teilnehmen wollen? Welche Gründe spielen bei
47 der Entscheidung eine Rolle?*
- 48 Ich hätte mir mehr Anleitung/ Begleitung durch die Studierenden/ Wissenschaftler
49 gewünscht. Beispielseinheiten in regelmäßigen Abständen und zudem gab es
50 erschwerte Bedingungen durch hohe Fluktuation.
- 51 *Wie wurden die Interventionen von den Kindern angenommen? (Hat den Kindern es
52 Spaß gemacht oder waren sie abgeneigt?)*
- 53 Bewegungsfreudige Kinder haben das Angebot gut und gerne angenommen. Andere
54 Schüler*innen jedoch mussten extrem stark animiert und motiviert werden.
- 55 *Welche Veränderung haben die Eltern an Ihren Kindern bemerkt?*
- 56 Keine Rückmeldung.
- 57
- 58 Bewegungsklasse Recklinghausen
- 59 *Welchen Effekt haben Sie durch die jeweiligen Interventionen gespürt?*
- 60 In der Bewegungsklassen waren die Kinder schnell motiviert, gewöhnten sich schnell
61 an die Bewegungen und hatten viel Spaß dabei. Nach den Bewegungsphasen waren
62 sie viel motivierter und konzentrierter bei den Arbeiten.
- 63 *Hat sich etwas im Klassengefüge verändert? Wenn Ja, was hat sich verändert?*

- 64 Sie waren alle sehr stolz Sportklasse zu sein und haben sich alle sehr angestrengt,
65 die Übungen zu erfüllen. Der Zusammenhalt hat sich eher verstärkt.
- 66 *Welche Intervention wurde durchgeführt und in welchem Maße?*
- 67 Wir haben jeden Tag zwei Mal täglich sportliche Übungen durchgeführt.
- 68 *Gab es ein Feedback von den Eltern? Wenn Ja, welches Feedback?*
- 69 Begeisterung auf allen Seiten.
- 70 *Würden Sie an der Studie noch mal teilnehmen wollen? Welche Gründe spielen bei*
71 *der Entscheidung eine Rolle?*
- 72 Ja, aber mit festen Terminen zur Überprüfung.
- 73 *Wie wurden die Interventionen von den Kindern angenommen? (Hat den Kindern es*
74 *Spaß gemacht oder waren sie abgeneigt?)*
- 75 Hat allen Spaß gemacht.
- 76

Schriftliche Befragung

Allgemein

1. *Klassenart*

Bewegungs-klasse

Leseklasse

2. *Altersstufe*

Ü6 Kinder

Ü9 Kinder

Effekt

1. Welchen Effekt haben Sie durch die jeweiligen Interventionen gespürt?

2. Hat sich etwas im Klassengefüge verändert? Wenn Ja, was hat sich verändert?

Reaktion und Umsetzung

1. Welche Intervention wurde durchgeführt und in welchem Maße?

2. Gab es ein Feedback von den Eltern? Wenn Ja, welches Feedback?

3. Würden Sie an der Studie noch mal teilnehmen wollen? Welche Gründe spielen bei der Entscheidung eine Rolle?

Verhalten der Kinder

1. Wie wurden die Interventionen von den Kindern angenommen? (Hat den Kindern es Spaß gemacht oder waren sie abgeneigt?)

2. Welche Veränderung haben die Eltern an Ihren Kindern bemerkt?

Anhang C

1 **Gedächtnisprotokoll_10.11.2021_Gelsenkirchen_1. Erhebung**

2 Wir (Studierende) und Rabea kamen an der Schule in der Bergmannsglückstraße in
3 Gelsenkirchen an und erkundeten uns nach passenden Räumlichkeiten. Zunächst war
4 nicht direkt ersichtlich welche Räumlichkeit sich am besten anbieten würde und wur-
5 den schlussendlich in einen Klassenraum geschickt. Der Raum war klein, weshalb wir
6 mit einigen Stationen auf den Flur ausweichen mussten. Die Lehrer*innen in den ne-
7 benanstehenden Klassenräumen, bekundeten ihren Unmut vor und auch während der
8 Testungen der Kinder. Vor allem die Lehrerin einer Intensivklasse, welche sich unmit-
9 telbar daneben befand, äußerte sich vermehrt negativ und behinderte die Testungen
10 durch rücksichtsloses Durchqueren der Testsituationen. Wir führten die Testungen
11 dennoch weiter durch, da keine weiteren Räumlichkeiten zur Verfügung standen.

12 Durch einen Konflikt in der Pause, welcher laut Aussagen der Kinder in Handgreiflich-
13 keiten endete, erwies sich die weitere Durchführung von Testungen als schwierig. Ei-
14 nige der Kinder waren sehr aufgebracht und ließen sich nur schwer runterbringen. Zu-
15 nächst waren einige der Studierenden überfordert mit dieser Situation, da noch keine
16 vorherigen Berührungspunkte mit Kindern aus diesem Förderschwerpunkt erfolgte und
17 zusätzlich die gesamte Testsituation neu war.

18 **Gedächtnisprotokoll_12.11.2021_Bottrop_1. Erhebung**

19 Die Erhebungen verliefen unproblematisch und auch die Räumlichkeiten waren für die
20 Testungen angemessen groß. Auffallend war hierbei, dass einige der Kinder eine enge
21 Betreuung durch Integrationshelfer*innen benötigten und dieser somit unterstützend
22 an der Erhebung teilnahmen.

23 **Gedächtnisprotokoll_17.11.2021_Recklinghausen_1. Erhebung**

24 In Recklinghausen wurde uns direkt zu Beginn die Aula zur Verfügung gestellt. Auf-
25 grund des Alters der Kinder, da es sich hierbei auch um eine Grundschule handelte,
26 waren die Kinder im Gegensatz zu den Kindern der vorherigen Erhebungen umgäng-
27 licher und aufgeschlossener gegenüber den Aufgaben. Jedoch verhielten sich einige
28 Kinder in der konkreten Testsituation schüchtern und absolvierten die einzelne Auf-
29 gabe sehr ruhig. Da innerhalb dieser Schule, alle Klassen am selben Standort lagen,
30 war die Erhebung durch die größere Menge an Schüler*innen deutlich länger als zuvor.
31 Durch ihren Förderschwerpunkt „Sprache“ kam es vereinzelt zu Missverständnissen
32 bei der Erklärung und Absolvierung der Aufgaben.

33 **Gedächtnisprotokoll_20.01.2022_Gelsenkirchen_2. Erhebung**

34 Bei der zweiten Erhebung in Gelsenkirchen durften wir diesmal direkt in die Aula wel-
35 che groß genug war, um alle Stationen angemessen aufbauen zu können. Allerdings
36 kamen durch diese Räumlichkeiten neue Störquellen zum Vorschein. Zum einen merk-
37 ten mehrere Kinder an, dass der Boden rutschig sei und somit Aufgaben wie beispiels-
38 weise balancieren oder die Liniensprünge schwieriger zu absolvieren waren. Außer-
39 dem stellte sich die Pausenklingel als weitere Störung dar. Diese war außerordentlich
40 laut und sowohl die Kinder als auch die Studierenden erschrecken und die Testung
41 musste zeitweise aufgrund dessen unterbrochen werden.

42 Durch eigenständige Aussagen fanden wir heraus, dass sich viele Kinder an die Auf-
43 gaben aus der ersten Erhebungsphase erinnern konnten und diese somit laut eigenen
44 Aussagen einfacher waren. Ebenso hatten Kinder vereinzelt die Sorge, dass die Er-
45 gebnisse der Testungen in ihre Noten miteinfließen („Fließt das hier in meine Note mit
46 ein?“).

47 Auch fielen einige Kommentare bezüglich der Bewegungspausen. Die Kinder berich-
48 teten, dass diese Pausen nicht regelmäßig gemacht werden und sie somit nicht ver-
49 stehen, wieso wir dann die Testungen mit ihnen machen.

50 **Gedächtnisprotokoll_21.01.2021_Bottrop_2. Erhebung**

51 Die zweite Erhebung in Bottrop war die erste Erhebung, die ohne die Begleitung von
52 Rabea Wienholt stattfand. Es wurde die gleiche Räumlichkeit wie zuvor genutzt, wes-
53 halb der Aufbau und die Durchführung ohne Probleme verlief.

54 Dennoch fielen vereinzelt Aussagen wie, „Ich möchte das nicht machen, die anderen
55 Kinder gucken zu“ und „Ich kann das nicht“. Hierdurch wurde deutlich, dass die Kinder
56 in den Testsituationen unwohl fühlten.

57 **Gedächtnisprotokoll_15.05.2022_Gelsenkirchen_3. Erhebung**

58 Im Gegensatz zu den anderen Erhebungen wurden bei der dritten Erhebung weniger
59 Kinder aufgrund einer hohen Fluktuation vor allem innerhalb der Bewegungsklasse
60 erhoben. Diese Schüler*innen konnten aufgrund von folglich fehlenden Einverständ-
61 niserklärungen nicht erhoben werden.

62 An diesem Tag wurde eine schriftliche Befragung der Lehrer*innen in der Bewegungs-
63 sowie Leseklasse durchgeführt. Nach der Erhebung wurden diese von uns analysiert
64 und zusammen besprochen.

65 **Gedächtnisprotokoll_19.05.2022_Bottrop_3. Erhebung**

66 Auffallend an dieser Erhebung war, dass lediglich drei Studentinnen, aufgrund von
67 Krankheitsfällen, die Testungen durchführen mussten und es somit länger dauerte.

11. Eidesstaatliche Versicherung

Wir versichern, dass wir diese Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich und sinngemäß übernommenen Textstellen als solche kenntlich gemacht haben. Dies gilt auch für die in der

Arbeit enthaltenen Zeichnungen, Skizzen und graphischen Darstellungen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Dortmund, 01.07.2022